How Costly are Trading Biases?*

Hee-Seo Han†
Xindi He‡
Daniel Weagley§

August 28, 2025

Abstract

We analyze a large set of stock selection trading biases that have been identified in top finance journal articles over the past 75 years. Using both retail and institutional trade-level data over 14.5 years, we assess the prevalence of each bias, whether the biases are associated with lower (or higher) future returns, and how their prevalence and performance depends on market conditions. We find that around 20 percent of studied biases are not more prevalent in our sample of retail investors than would be expected from random trading, and retail investors exhibit three times as many biases as institutions. Examining retail investor trades, we find that the average trading bias is associated with a 128-basis-point lower return over the next 12 months and that this underperformance is mainly a within-investor phenomenon. We find that the prevalence of biases varies with market conditions, and that the underperformance related to the average bias is especially severe in "bad times." Our results shed light on the nature of investor trading biases and their costs.

Keywords: investor biases, stock selection, retail investors, heuristics, financial decision-making, behavioral finance.
JEL Classification: G11, G12, G40, G41.

*An older version of this paper was circulated under the title "How Costly are Trading Heuristics?" We are grateful to Andy Puckett, Dan Egan, Mete Kilic, Jing Zhao, Joshy Jacob, and seminar and conference participants at the Boulder Summer Conference on Consumer Financial Decision Making, International Behavioural Finance Conference, UConn Finance Conference, China International Conference in Finance, Georgia Institute of Technology, and University of Tennessee for helpful comments.

†Paul Merage School of Business, University of California at Irvine; email: heesh1@uci.edu.
‡Scheller College of Business, Georgia Institute of Technology; email: xindi.he@scheller.gatech.edu.
§University of Tennessee; email: dweagley@utk.edu.

1 Introduction

Individual investors are an integral set of traders in modern stock markets, with evidence that they influence both market liquidity (e.g., Kaniel, Saar, and Titman, 2008; Eaton, Green, Roseman, and Wu, 2022) and asset prices (e.g., Kumar and Lee, 2006; Barber, Odean, and Zhu, 2008; Davis, Knüpfer, Soerlie Kvaerner, Sen Dogan, and Vokata, 2025). There is also significant evidence that they exhibit systematic trading behaviors (i.e., trading biases). These trading biases can manifest due to behavioral biases in information processing or decision making, non-standard preferences, limited attention, or the use of heuristics to make investment decisions. Trading biases could also arise due to relative advantages, for instance, in information, tax treatment, or liquidity provision. Depending on the underlying driver(s), trades associated with these biases may be associated with lower or higher relative future returns. In this paper, we synthesize the literature on stock selection-related trading biases identified over the past 75 years and holistically study their prevalence and associated performance using 14.5 years of trade-level data. By studying a large set of biases jointly over a relatively long time-horizon using a novel methodology, we uncover new insights into the investment decision-making process.

The main takeaways of our paper are that trading biases are common among retail investor stock purchases, biases are costly on average, and there is significant time variation in both the prevalence and cost of retail investor trading biases. Highlights of our results include: (1) the finance profession has identified a significant number of stock selection biases of which we are able to study 22 biases, (2) around 20% of previously documented trading biases are not more prevalent in our sample of retail investors than would be expected by random trading, (3) retail investors systematically exhibit a much larger number of stock selection biases (18 biases) than institutional investors (6 biases), (4) the prevalence of many trading biases varies with market conditions, (5) each additional bias associated with a retail stock purchase is associated with 128 bps lower returns over the next 12 months, (6) in periods

of high volatility or low market returns, bias exposure is greater and the underperformance is more severe. For instance, a one standard deviation increase in market volatility decreases the future 12-month return associated with an additional bias by about 73% (105 bps) on average. Furthermore, we find that (7) there is significant performance variation between biases with some biases actually associated with better performance, and (8) the prevalence of these biases differs between investors and is related to their characteristics.

Although it is commonly understood that retail investors tend to underperform and suffer from trading biases, our study contributes to the literature on investor behavior in three important ways. First, we more accurately assess the performance associated with a large set of biases and disentangle the drivers of bias-related performance. Our approach allows us to test whether bias-related performance captures trade-invariant investor skill, market timing, stock characteristics, or relative stock selection ability conditional on these factors. Second, we challenge the treatment of trading biases as time-invariant characteristics of individual investor trades or the interpretation of results in such a manner.[1] We show that the prevalence and costs associated biases vary with market conditions. Our results underscore the importance of studying investor behavior over a longer time horizon and in different market conditions, and can explain why studies conducted during different time periods can arrive at different results (e.g., retail investors appear contrarian in some studies and as momentum traders or extrapolators in other studies). Third, and relatedly, our results also highlight the importance of an explicit counterfactual in studying investor behavior to ensure that the documented prevalence of a behavior is significant and to help tease out the drivers of bias-related performance.

We begin our analysis by collecting a large set of previously identified trading biases. To do so, we process, using AI and human research assistants, all papers published in top-tier finance journals over the past 75 years. Academics in the finance literature have uncovered

[1]Much of the prior literature studies the average returns of individual investors or the performance of stocks with characteristics associated with a particular bias (e.g., returns on "lottery-like stocks").

a significant number of retail investor stock selection biases over the past 75 years with the vast majority identified in the late 1990s and 2000s. For our purposes, we identify dozens of papers covering 22 trading biases related to an investor's stock purchase decision that are feasible for us to study in our setting. This is a subset of all stock selection biases as some we cannot study in our setting and others may have only been tested in non-top-tier finance journals. Using transaction-level data on individual investors' trading activity in the Finnish market over 14.5 years, from January 1995 through June 2009, we find that at least 90.51% of retail investor's trades are consistent with one bias and the vast majority are consistent with multiple biases.[2]

The estimated high bias rate is partially driven by false positives. We address this issue by comparing the bias rate in our real trading data with the results from simulated trading data based on random stock selection. The simulated data provide an important counterfactual and help ensure that the outcomes we examine are not due to random chance. For retail investors, we find that approximately one-fifth of the biases we study are observed less often than would be expected with random trading, suggesting that they are not systematic biases among our sample of investors. Individual investors do exhibit 18 biases, which is the set on which we focus most of our analysis. For institutional investors, they exhibit only six biases, with around three-quarters of the trading biases found less often than would be expected by random chance. These results suggest that when investors have more resources and expertise, they exhibit fewer trading biases.

We next assess the performance costs associated with trading biases. We focus our analysis on the future returns on selected stocks. Whether we expect a trading bias to be associated with poor performance depends on the underlying driver(s) of the bias and the degree of market efficiency. For example, limited attention-driven trading biases could lead investors toward overpriced stocks or to push up prices of stocks that underperform in the

[2]Like most studies of trading biases, we do not observe the actual decision-making process of investors. Instead, we infer the investment process from the observable features of the transaction. We use bias-specific conditions to identify each trade as consistent with or not with the bias.

future (Barber and Odean, 2008), trading biases derived from an informational advantage could lead to better performance, and the trading biases that derive from the use of heuristics could lead to worse or better performance depending on the heuristic model.[3]

We find that unconditionally there is significant underperformance associated with the average trading bias in the longer run. For each additional bias linked to a trade, the average 12-month return on the selected stock is 128 basis points lower (t-statistic of -7.58). Although we find that the relative return is near zero after 3 months, the returns steadily decline up to two years after the transaction. This low performance could be due to the biased stock selection methodology leading to poor stock selection, potentially due to simplified decision making. Beyond the potential role of poor stock selection, the lower associated returns could also be due to multiple other factors including: (1) trading biases are associated with particular stock characteristics that are related to lower future returns, (2) investors who exhibit a greater number of biases may have lower levels of skill, (3) and poor timing of trades associated with more biases, wherein investor trades are more biased at times when future returns are likely to be lower. We provide additional tests to shed light on the particular channels.

To examine the role of stock characteristics in the underperformance of biased trades, we control for stock systematic risk and compare the relative performance of bias-related trades in the real data to the relative performance of bias-related trades in the simulated data. We find little evidence that differences in stock characteristics can explain the relative underperformance. To control for the effect of investor skill, we include investor fixed effects and trade characteristics such as the principal value of a trade and the number of years from the investor's first trade to the current trade. We find these have little impact on the relative performance estimate. These results suggest that trading biases are not just capturing variation in skill across investors; instead, the relationship is mainly driven by

[3]The use of heuristics in complex environments such as asset markets has been shown to lead to better performance(Gigerenzer, 2008; Gigerenzer and Brighton, 2009).

within-investor variation in trading biases.

We next examine whether market timing can explain the relative underperformance of bias-related trades. We find that including date fixed effects does not significantly impact our return estimates. We also find that using market-adjusted returns, instead of raw returns, does not materially impact our relative performance estimates. These latter results suggest that market timing plays some role in the underperformance of bias-related trades, but that the majority of the underperformance is likely due to poor stock selection with more biased trades.

We find similar results when we examine the performance of bias-related trades made by institutional investors. We first examine the same set of 18 biases we examine for retail investors. We find a negative relationship between the number of biases associated with a trade and future returns of selected stocks, though the coefficients are statistically weaker than those for retail investor trades. The results weaken further when we focus on the six biases exhibited by institutional investors. The institutional investor results suggest that, while institutional biases are costly in our sample, they avoid much of the costs by not exhibiting as many biases.

Returning to retail investor trades, we further examine the impact on performance at the individual bias level in a pooled regression. In the baseline specification, we find that seven of the 18 biases are negatively related to performance at the 12-month horizon, and five are positively related to performance. We find similar results — eight related to underperformance and four related to outperformance — when adjusting for trade characteristics, market beta, and investor skill. When we additionally compare return outcomes for biased trades relative to randomly selected stock trades that appear to be biased (our method for conditioning on stock characteristics), we find eight biases exhibit relative underperformance, four exhibit outperformance, and six do not exhibit any differential in performance. Overall, these results highlight the heterogeneity in the costs associated with

particular biases and the nature of these costs.

After establishing the baseline performance results, we turn our attention to time-variation in the costs and prevalence of biases. Examining all biases jointly, we test whether bias-related performance depends on market conditions. We find that the underperformance associated with trading biases is mainly concentrated in periods of high past market volatility and low past market returns. We find that a one standard-deviation increase in past market volatility decreases the 12-month return associated with a trading bias by about 1.42 pps. These results suggest that bias-related trades are not necessarily costly during good states of the world, where they are associated with neutral or even positive performance, but bear significant costs during downturns, when market tends to be more volatile and noisier, and investors' ability to detect the cost is diminished.

The prevalence of biases also varies with market conditions. We find a greater number of biases in "bad times," when the dividend yield is higher, market returns are lower, and volatility is greater. These are precisely the times when the costs associated with biases are greatest, suggesting they become more biased when it hurts the most.

After documenting the costs of trading biases, we explore the determinants. We first provide a simple model for the decision to use a particular trading strategy. Our model incorporates potential utility for investors who use a biased trading method either because of lower information processing costs or because of some utility derived from the bias exposure. The investor also learns about the impact of the trading method on return performance by observing the returns and updating beliefs through Bayesian updating. The model predicts that investors will be more likely to continue to use a biased trading method if (1) the exposure to the trading bias is not as detrimental to their performance, (2) they derive greater utility from its use, or (3) the benchmark (i.e., market) return is noisier. These observations, in combination with our empirical findings, lead to an explanation for the prevalence of biased stock selection models from a learning perspective. Investors keep

trading with biases because it is very difficult for them to learn about the associated cost. During good times, the cost of biases is low, so they do update beliefs significantly and won't quit a given bias as it brings them utility or convenience. During bad times, the cost of biases is high, but the market also becomes noisier, which prevents them from sufficiently learning about the cost.

Our last set of tests examine which types of traders exhibit more trading biases. We find that females and older investors are slightly more likely to exhibit trading biases in general. Perhaps surprisingly, we find that investors with more years of experience are more likely to exhibit trading biases. We also interact the demographic variables with market conditions and find that more experienced investors, even though displaying more biases in their trading during good times (i.e., high market return and low market volatility), reduce their trading's exposure to biases during bad times.

Our results contribute to the literature on the trading behavior and performance of retail investors. Several articles show how retail investors' trading decisions are influenced by behavioral biases, such as the rank effect (Hartzmark, 2015), the salient stock bias (Barber, Huang, Odean, and Schwarz, 2022), and the familiarity bias (Keloharju, Knüpfer, and Linnainmaa, 2012). Our work focuses on the trading biases observed in retail investor trades, not the potential underlying behavioral biases that may lead to the trading biases. We are relatively agnostic on the underlying driver(s) of any particular trading bias in our setting.

Previous studies attribute differences in retail investor performance to variations in investor characteristics such as skill (Coval, Hirshleifer, and Shumway, 2021), cognitive ability (Korniotis and Kumar, 2011; Grinblatt, Keloharju, and Linnainmaa, 2012), market timing ability (Keppo, Shumway, and Weagley, 2021), overconfidence (Barber and Odean, 2000, 2001; Grinblatt and Keloharju, 2009), and trading experience (Seru, Shumway, and Stoffman, 2010). Recent work examining aggregated holdings by investors using the Robinhood

brokerage platform provides mixed evidence on the performance of stocks held by retail investors (Barber et al., 2022; Welch, 2022). We extend this literature by showing the costs associated with biased trades and how these costs and the prevalence of these biases vary with market conditions.

2 Trading Bias Selection

Our goal is to identify and analyze the stock purchase trading biases of individual investors. In this section, we detail our process for identifying the set of trading biases to study and how we identify the exposure to these trading biases in our trade-level data.

To identify our set of trading biases, we first collect all articles published in the top three finance journals in the last 79 years — *Journal of Finance* (JF), *Journal of Financial Economics* (JFE), and *Review of Financial Studies* (RFS) and use artificial intelligence to identify papers that discuss stock selection trading biases of retail investors. We construct and test the prompt such that it reduces the false negative omission. Then, we read through all the selected papers to remove false positive selections.

2.1 Paper Collection

We use a systematic data collection approach to identify articles studying trading biases used for stock trading. We manually download all articles from the official websites of the top three finance journals and exclude non-research content, such as announcements, discussions, cover pages, and other ancillary materials. The publication time range and the corresponding number of articles for each journal are summarized in Table 1.

2.2 Classification Method

After downloading all the articles, we use the ChatGPT-4 model through an API to classify the papers. Our goal is to identify all research that potentially documents a trading bias. Our methodology is designed to optimize the trade-offs between Type II error (false negative) rate, API usage costs, and computational runtime. We develop a two-step classification method. Since trading bias papers are a subset of behavioral finance papers, we first task GPT-4 with identifying behavioral finance papers from all papers and then detect the trading bias papers from the set of behavioral finance papers. This way, we can avoid incurring unnecessarily high cost from the first step with massive paper reading by feeding only a small portion of each paper, since behavioral finance determination is relatively easier; we can also improve the accuracy in the second step, where we start from a much smaller set of papers, by feeding in the full text of the papers, since trading bias determination involves more complicated criteria. We intentionally design the classification process to be inclusive, minimizing the Type II error rate in both steps to ensure that we capture nearly all papers related to trading biases.

Step 1: Behavioral finance paper detection

To identify behavioral finance papers, we analyze the title, abstract, and conclusion of each paper. For older papers lacking clear structure or machine-readable formats, Optical Character Recognition (OCR) is applied. Papers that exceed 10 pages are processed by extracting the first eighth and last eighth of the text after excluding references, while shorter papers use the entire text. After removing non-textual elements such as plots, special characters, and formulas, the average token count of the input text is approximately 4,800.

The classification process uses the following prompt:

"The text I provided to you is the title, abstract and conclusion of a finance academic paper. If you think the paper is not related to behavioral finance, return 0. If it is a paper related to behavioral finance, then return 1. Just return numbers to me; do not return

anything else."

Step 2: Trading bias paper detection

The second step classifies biase-related articles from the subset of behavioral finance papers identified in Step 1. To maximize accuracy, the full text of each paper is analyzed. Similarly to Step 1, OCR is employed for older papers, with text extraction applied to all pages except reference pages. After preprocessing, including the removal of non-text elements, the average token count of the input text is approximately 95,000. We provide the ChatGPT-4 model with the following prompt:

"You are a world-class finance academic in a top finance department. Read the behavioral finance paper and decide whether it documents a bias in the stock purchase decision using trade-level data. Presence of generic terms like "stock," "equity," or "return" is not enough, and studies that only examine asset-pricing, corporate finance, macro factors, statistical biases, or other asset classes without analyzing equity trades do not count. Output 1 if the paper documents a bias in stock purchases and 0 otherwise. Return only the single digit without any additional text. Do not explain your decision."

2.3 ChatGPT Classification Results

The large number of potential articles (646) remaining after Step 2 suggests that the prompt and process were very inclusive. The classification results are shown in Table 1. The yearly trends for the number of all published articles, behavioral finance articles, and trading bias related articles in the top three journals are shown in Figure 1. For all three journals, we can see an overall increasing trend in the number of papers potentially related to trading biases over the last 20 years.

Many of the 646 identified papers are "false positive" results due to the intentional inclusiveness of our methodology. We further narrowed the set of papers down through a

subsequent manual review as discussed in the next section.

2.3.1 Robustness Check

To ensure the reliability and accuracy of our classifications, we performed a bootstrap validation experiment to repeat the two-step selection procedure using different subsets of papers identified as related to trading biases in our manual review. The primary focus of this analysis is to examine the Type II error rate of the selection procedure. We construct a set of tests and apply the bootstrap methodology to evaluate the Type II error rate of our classification model.

The set of test papers is constructed to preserve the ratio of papers between the three journals: *Journal of Finance* (JF), *Journal of Financial Economics* (JFE), and *Review of Financial Studies* (RFS) at approximately 5:3.5:2.5. Additionally, the year-by-year distribution within each journal also mirrors that of the original full set of papers. The test set comprises 100 papers in total, representing a balanced sample for robust analysis.

We ran 100 iterations with 10 papers randomly selected in each round. We pose the same questions to the ChatGPT-4 model as in our main selection process. Different random seeds are utilized to ensure that each paper appears at least twice during the entire bootstrap procedure. This approach provides comprehensive coverage of the data set. The balanced nature of the bootstrap procedure is visually presented in Figure A1.

The results of the 100-round bootstrap experiment are summarized in Table 2. We find that the Type II error rate for the first step of the procedure is 0. While the rate for the second step is non-zero, we consider the selection methodology to be robust and reliable given the complicated nature of the task.

2.3.2 Manual Review

The final step is to process the 646 articles potentially related to trading biases identified by the ChatGPT-4 model. We manually review each paper to determine whether it is related to biases.

We exclude biases that are infeasible to examine due to data limitations. We ultimately have a list of 37 papers related to stock purchase biases that we are able to examine in our setting, which corresponds to twenty-two identified biases. We summarize the set of biases in Table A1 and discuss the identified biases in detail below.

2.4 The Final Set of Stock Trading Biases and How We Identify Plausibly Biased Trades in the Data

Our analysis of the literature reveals a wide range of trading biases for stock purchases. We give each bias a name based on the nature of the bias and the associated papers. We either take the definition directly from existing papers or slightly modify the original definition to fit our data and setting. In particular, each bias is a bias toward a particular stock characteristic or combination of stock and investor characteristics. We identify the characteristics or conditions that a transaction must have to be in line with the bias of interest. We attempt to choose simple conditions that are most in line with the overall nature of the bias.

In the following, we describe the final set of 22 trading biases in alphabetical order, the literature on which each bias is based, and the conditions we use to identify the possible exposure to biases in the trading data. We provide a tabular summary of this information in the Appendix Table A1.

- *Abnormal Trading Volume* is based on Barber and Odean (2008), Da, Engelberg, and Gao (2011) and Gargano and Rossi (2018). Barber and Odean (2008) find that

individual investors tend to buy attention-grabbing stocks with a higher abnormal trading volume. Da et al. (2011) introduce Google search volume index as a measure of investor attention and document its relationship to stock price movements. Gargano and Rossi (2018) investigate which individual investors are the most attentive, how investors allocate their attention, and the relation between investor attention and performance. We define *Abnormal Trading Volume* as an indicator that equals 1 if the purchased stock has an abnormal trading volume (Barber and Odean, 2008) above the median.

- *Chase Dividend* is based on Rantapuska (2008), who document that households are more likely to buy a stock before its ex-dividend day than after, especially for high-dividend stocks. We define *Chase Dividend* as an indicator that equals 1 if the stock is purchased on a day that falls between the day of the announcement of the dividend and the record date of the dividend claim.

- *Contrarian* is inspired by Lakonishok, Shleifer, and Vishny (1994), Grinblatt and Keloharju (2001b), Kaniel et al. (2008). Kaniel et al. (2008) find that individuals tend to be contrarians as they buy stocks following declines in the previous month and sell following price increases; and Lakonishok et al. (1994) provide evidence that value strategies generate higher returns. We define *Contrarian* as an indicator that equals 1 if the stock purchased has a past 20-day return in the bottom 10 percentile.

- *Easy Name* is based on Green and Jame (2013), who find that companies with short and easy-to-pronounce names have greater ownership breadth, greater share turnover, lower transaction price impacts, and higher valuation ratios. We define *Easy Name* as an indicator that equals 1 if the stock name consists of a single word.

- *Holding Period* is based on Massa and Simonov (2006), who find households invest in stocks that they are more familiar with through a long holding period. We define

Holding Period as an indicator that equals 1 if the investor buys a stock they have held for at least six months.

- *High Dividend* is based on Graham and Kumar (2006), Hartzmark and Solomon (2019), Daniel, Garlappi, and Xiao (2021). Graham and Kumar (2006) document that low-income investors disproportionately purchase stocks following dividend announcements but before the ex-dividend day. Hartzmark and Solomon (2019) also find that many individual investors, mutual funds, and institutions trade as if dividends and capital gains are disconnected attributes, not fully understanding that dividends result in price drops. We define *High Dividend* as an indicator that equals 1 if the purchased stock falls into the top decile of all stocks sorted on prior year's dividends.

- *Top Industry Experience* is based on Kaustia and Knüpfer (2008), Huang (2019) and Malmendier (2021). Malmendier (2021) documents the long-lasting effect of past experiences on beliefs and risk taking, Huang (2019) examines the effect of prior investment experience in specific industries on subsequent investment decisions, and Kaustia and Knüpfer (2008) analyze the role of past returns on future investment behaviors, highlighting the influence of individual experience on risk-taking. We define *Industry Investment Experience* as an indicator that equals 1 if a stock purchased is not previously owned and in an industry where the investor experienced the top 10 percentile market-adjusted return in the previous year.

- *Integer Price*: based on Kuo, Lin, and Zhao (2015), who find that investors with higher limit order submission ratios at round numbers suffer greater losses in their round-numbered and non-round-numbered limit orders, market orders, and round-trip trades. We define *Integer Price* as an indicator that equals 1 if the transaction price is within 5 cents distance to an integer price.

- *Local Company* is based on Huberman (2001), Grinblatt and Keloharju (2001a), Feng and Seasholes (2004), Ivković and Weisbenner (2005), Ivković and Weisbenner (2005)

, and Seasholes and Zhu (2010). Huberman (2001) highlights the familiarity bias in investment choices. Feng and Seasholes (2004) find that the purchases and sales of stocks are highly geographically correlated. Ivković and Weisbenner (2005) find that households show a strong preference for local investments. Seasholes and Zhu (2010) investigate the geographical clustering of retail investors and how it relates to performance. We define *Local Company* as an indicator that equals 1 if the purchased stock has the same municipality as that of the investor.

- *Lottery-Like Stocks* is based on Kumar (2009) and Eraker and Ready (2015), who find individual investors prefer stocks with lottery features at the aggregate level and with positively skewed payoffs. We define *Lottery-Like Stocks* as an indicator that equals 1 if the stock purchased is in the highest 50th percentile of previous six month idiosyncratic volatility, and in the highest 50th percentile of previous six month idiosyncratic skewness, and in the lowest 50th percentile of stock price.

- *Low Price* is based on Green and Hwang (2009) and Birru and Wang (2016). Birru and Wang (2016) explore the psychology of stock price levels and provide evidence that investors suffer from a nominal price illusion in which they overestimate the room to grow for low-priced stocks relative to high-priced stocks. Green and Hwang (2009) document nominal price-based comovement. We define *Low Price* as an indicator that equals 1 if the purchased stock falls into the bottom decile of all stocks sorted according to the nominal price the previous day in the same currency (Euro).

- *Max Return* is based on Bali, Cakici, and Whitelaw (2011), who investigate the significance of extreme positive returns in the cross-sectional pricing of stocks. Liu, Wang, Yu, and Zhao (2020) also uses the extreme positive returns to measure investors' preference. We define *Max Return* as an indicator that equals 1 if the stock purchased falls into the top decile of all stocks sorted on their last month's maximum daily returns.

- *Momentum* is inspired by Grinblatt and Keloharju (2000) and Liao, Peng, and Zhu

(2022), who suggest the extrapolative beliefs about stock returns among retail investors. We define *Momentum* as an indicator equals 1 if the stock purchased has past 20-day return in the top 10 percentile.

- *Naïve Buying Diversification (NBD)* is from Gathergood, Hirshleifer, Leake, Sakaguchi, and Stewart (2023), who show that individual investors equally balance the values of same-day purchases of multiple assets, i.e., naïve buying diversification. We define *NBD* as an indicator equals 1 if multiple stocks are purchased on the same day with the same amount of money up to a difference of 5%.

- *Peer Influence Industry* and *Peer Influence Industry* are inspired by Ivković and Weisbenner (2007), Hvide and Östberg (2015), Huang, Hwang, and Lou (2021), and Escobar and Pedraza (2023). Escobar and Pedraza (2023) study the influence of social interactions on equity trading and find that investors react more to peer gains than to peer losses. Hvide and Östberg (2015) find that the stock investment decisions of individuals are positively correlated with those of co-workers. Ivković and Weisbenner (2007) study the relation between household stock purchases and stock purchases made by their neighbors. Huang et al. (2021) investigate the peer effect at the industry level. We define *Peer Influence Industry* (*Peer Influence Industry*) as an indicator that equals 1 if the ratio of the fraction of investors in the same zip code who purchased a stock in the same industry (the same stock) to the fraction of all other investors who purchased stocks in the same industry (the same stock) is the top industry (the top 10% of stocks) within that zip-code month, and the ratio is at least 1.

- *Return Sorting Top* and *Return Sorting Bottom* are based on Grinblatt and Keloharju (2001b), Da, Huang, and Jin (2021) and Jegadeesh and Titman (1993), where Grinblatt and Keloharju (2001b) find that past returns and historical price patterns, such as being at a monthly high or low, affect trading; Da et al. (2021) find that investors extrapolate from stocks recent past returns, with more weight on more recent returns;

and Jegadeesh and Titman (1993) introduces momentum investing as a strategy based on the persistence of stock performance. We define *Return Sorting Top* (*Return Sorting Bottom*) as an indicator that equals to 1 if the stock purchased falls into top (bottom) 10 percentile of all stocks sorted on yesterday's returns.

- *Salient Returns* is based on Cosemans and Frehen (2021), who find that Retail investors overweight salient past returns. We define *Salient Returns* as an indicator that equals 1 if the purchased stock is in the top 10 percentile of salient returns (measured following Cosemans and Frehen (2021)) over the past month.

- *Small Cap* is inspired by Kumar and Lee (2006), who find that retail investors are highly concentrated in small-cap stocks. We define *Small Cap* as an indicator that equals 1 if the purchased stock is in the top 10 percentile of all stocks sorted on prior day's market capitalization.

- *Value* is inspired by Kumar and Lee (2006), who find that retail investors are highly concentrated in value stocks. We define *Value* as an indicator that equals 1 if the purchased stock is in the top 10 percentile of all stocks sorted on last quarter's book-to-market ratio.

- *Volatility* is inspired by Dorn and Huberman (2010), who propose the preferred risk habitat hypothesis, showing that individual investors select stocks whose volatilities are commensurate with their risk aversion. We define *Volatility* as an indicator that equals 1 if the purchased stock has a daily realized volatility over the past year within 10% of the average realized volatility of the other stocks in their portfolio.

3 Data, Summary Statistics, and Simulations

We use data obtained from the Nordic Central Securities Depository (NCSD) on all transactions in Finnish securities from January 1995 to June 2009. Our main analyses focus

on the regular buy-side trades of retail investors during the sample period.[4]

In the original dataset, 1,386,356 individual investors were identified as having traded common stocks during the sample period. Of these, 591,254 conducted at least two trades, either buying or selling, during the sample period. From this group of investors, a random sample of 50,000 was selected for further analysis.

For some analysis, we examine institutional investors. We use their regular buy-side trades during the same period from the same dataset. There are 1,065 institutional investors.

The summary statistics for retail investors are presented in Table 3. On average, retail investors hold 7.72 stocks at the time of trade and remained active in the market for 5.48 years after their first trade.[5] The purchased stocks have an average market beta of 0.64 and an average trade size of 6032.82 euros, adjusted for inflation to the year 2000.

For brevity, we do not report the full summary statistics for the institutional investors. We find that on average institutional investors hold 85.32 stocks at the time of trade and remain active in the market for 4.85 years after their first trade. The traded stocks have an average market beta of 0.62 and an average trade size of 470,387.8 euros, adjusted for inflation to the year 2000.

3.1 Simulations

Our methodology identifies trades that are plausibly associated with a bias (e.g., a purchase of a stock that has performed well in the past is plausibly associated with a momentum bias). When studying biases, it is important to have a counterfactual to compare the rates of observed bias against to determine whether investors are actually biased in a particular manner. To address this issue, we propose a novel simulation methodology that

[4]Each trade has an explanation code; we focus on the regular buy-side trades with an explanation code equal to one.

[5]We observe holdings at the start of the sample, so for investors who began before 1995, we will under-estimate their experience.

uses simulated trading records as a benchmark. This allows us to determine whether the observed bias rate is beyond coincidence. We also use the performance of simulated trades as a benchmark in some tests of the performance consequences of biased trades.

Our simulation algorithm matches the timing and size (up to a range of 50%) of all trades in the real data. In other words, only the stock itself can be randomly chosen, while all other characteristics remain comparable between the simulated and real datasets. In both versions, we randomly draw the purchased stock with probabilities equal to the market weights of the stocks.

As discussed in more detail in Section 4, we compare the observed prevalence of biases in the actual data to those in the ten simulations with random seeds to gauge whether there is a greater prevalence than would be expected by random chance.

3.2 Prevalence of Stock Selection Biases

In Figure 2, we plot the distribution of the number of biases associated with each trade separately for the real data and the ten simulations. We focus only on the 18 biases that are more prevalent than what we find in the simulated data (as documented below in Section 4). A total of 90.51% of the trades are associated with multiple biases, and the median trade is associated with three biases, with significant variation in the number of associated biases between trades. The number of biases per trade is greater in the real data than in the simulated data (median of two). The distribution suggests that retail investor stock selection decisions are consistent with multiple biases.

4 Overall Prevalence of Individual Biases

We begin our analysis by examining the prevalence of each of the 22 trading biases. We benchmark the percentage of trades associated with each bias in our real trading data

against a similar percentage in the simulated trading data sets. The difference between the real and simulated frequency of bias captures the degree of bias above and beyond a simpler strategy of buying larger company stocks with greater probability. We view this difference as an indication of the actual degree of bias in our sample. We first examine relative prevalence for retail investors than for institutional investors.

4.1 Overall Prevalence of Biases Among Retail Investors

We begin by examining the prevalence of biases in our sample of retail investor trades. In Panel A of Figure 3, we compare the prevalence of each bias in our real trading data with the prevalence observed in simulated trading data. To ensure our benchmark is not affected by the choice of random seed, we run ten simulations of the trading data and plot the range, which tends to be small. The biases are ordered by the difference between their prevalence in real data and the maximum prevalence observed in the simulations.

Four biases are either not used more frequently or are used less frequently than in the simulated data, suggesting that they are not relevant biases in our sample or that investors are actively pursuing opposing strategies. Interestingly, investors tend to shy away from small capitalization stocks even more than our simulated data which puts higher probability on higher market capitalization stocks. In addition to *Small Cap*, the other biases that do not appear to be used more frequently than in the simulated data are: *Value*, *Low Price*, and *Lottery Like Stock*. The results suggest that these biases are not existent overall in our sample of investors. We exclude these four non-prevalent biases from the remainder of our analysis. Although *Lottery Like Stock* does not appear as a relevant bias in our sample, we do find that other related biases are relevant, such as *Salient Returns*, *Max Return*, and *Return Sorting Top*.

Overall, we find that 18 of the 22 biases are used more frequently in our real trading data than in any of the simulations. The biases that are more prevalent than in the simulated data,

in order of relative prevalence, are: *Abnormal Trading Volume*, *Contrarian*, *Peer Influence Industry*, *Peer Influence Stock*, *Volatility*, *Holding Period*, *Chase Dividend*, *Integer Price*, *Return Sorting Bottom*, *NBD*, *Easy Name*, *High Dividend*, *Local Company*, *Momentum*, *Top Industry Experience*, *Salient Returns*, *Max Return*, and *Return Sorting Top*.

4.2 Prevalence of Biases Among Institutional Investors

Next, we examine the biases of institutional investor stock purchases. Institutional investors provide an interesting contrast to our set of retail investors. Generally speaking, institutions have significantly more resources to make investment decisions. They are also subject to different incentives and constraints and may be less subject to behavioral bias than individual investors. Given these conditions, we would expect that institutions would be less likely to use biases for investment decision-making.

The prevalence results for institutional investors are presented in Panel B of Figure 3. We find that institutions only exhibit six (out of twenty-two) biases. The six biases are *Abnormal Trading Volume*, *Integer Price*, *Holding Period*, *NBD*, *Peer Influence Stock*, and *Contrarian*.

The existence of *Peer Influence Stock* and *Abnormal Trading Volume* is not necessarily surprising. The exposure to *Peer Influence* is consistent with Pool, Stoffman, and Yonker (2015) and can indicate information sharing between institutions located in the same area. The exposure to *Abnormal Trading Volume* could be due to trading with (Brunnermeier and Nagel, 2004) or against retail investors in stocks with a high volume of retail trading. Potentially, these biases could be beneficial for institutions, a prediction that we test in Section 5.1.1.

In the next section, we examine the performance costs associated with the exposure to prevalent biases used by retail and institutional investors.

5 The Relative Performance of Bias-related Trades

Investor trading biases can stem from several sources. They could be driven by departures from rational information processing, reliance on heuristics under cognitive constraints, or non-standard preferences. Biases may also reflect relative advantages between investors, for example, differences in access to information, tax considerations, or opportunities to provide liquidity. The relative performance of trades associated with a bias should depend on the nature of the bias and the context in which it is used. For example, the use of heuristics is a common explanation for several retail investor trading biases. Heuristics can simplify an investor's decision making, saving them time, effort, and other information collection costs. However, the use of heuristics can lead to suboptimal decisions (Tversky and Kahneman, 1974). In the context of trading, this could result in lower returns. On the other hand, recent work on ecological rationality argues that heuristics can lead to better decisions in some contexts (Gigerenzer, 2008), especially in complex environments with uncertainty (Neth and Gigerenzer, 2015). While uncovering the underlying driver(s) of each bias is beyond the scope of this paper, we are able to assess the performance associated with each bias and biases in general.

For this analysis, the outcome focus on is future returns because they are a first-order determinant of an investor's overall investment performance, are readily observable, and are easily quantifiable. We acknowledge that there may be other associated costs that are more difficult for us to estimate. Most of our analysis is on the performance of retail investor trades, with some additional analysis of institutional investor trades.

5.1 Biases and Future Returns

We begin our performance analysis by examining the relationship between the number of biases associated with a trade and the future performance of that trade. We begin by

examining only individual investor trades. We measure performance as the future return of the stock purchased over the next 12 months. We regress the future 12 month return of the purchased stock on the *Number of Biases* of the trade, which is the total number of biases associated with the trade.

In most specifications, we control for the purchased stock's market beta to capture variation in systematic risk across stocks, the size of the trade to control for the value relevance of the trade to the investor, and three time-varying investor characteristics (the logarithm of the number of current holdings, the logarithm of the investor's years of experience, and the logarithm of the investor's age) that capture time variation in the investor's activity and experience in the market. We also layer on various fixed effects when assessing potential explanations for the relationship between biases and future returns. We cluster standard errors at the investor and trade date levels.

In column 1 of Table 4, we establish the baseline relationship between bias exposure and future 12-month returns. We do not include controls or fixed effects in this baseline specification. We find that greater bias exposure is associated with lower future returns, with the average bias corresponding to a 170 basis-point lower return over the next 12 months, with a t-statistic of -7.71. Next, we explore potential explanations for the negative relationship between biases and future returns. Biases can be related to future returns through multiple potential channels. The first potential channel is that biases may be associated with particular stock characteristics that are related to future returns. For example, bias-related trades may involve securities with a lower systematic risk that should earn a lower return on average.

The second potential channel is that investors who exhibit biases may have generally lower levels of skill. The third channel is that biases may lead to suboptimal selection above and beyond any bias toward a particular stock characteristic. The fourth potential channel is timing, where investors are more biased at times when future returns are likely to be lower.

We next attempt to shed light on the various channels.

In column 2, we include controls for the stock market beta, trade size, and investor characteristics. The coefficient on *Number of Biases* is barely changed and more statistically significant, suggesting that the return relationship is not driven by differing levels of systematic risk, investor experience or holdings, or investors exhibiting more biases on less consequential trades.

In column 3, we further include investor fixed effects to control for time-invariant investor characteristics and skill. The inclusion of investor-fixed effects will allow us to estimate the performance of trades associated with the exposure to different numbers of biases within the same investor. We therefore ensure that we are not capturing time-invariant differences in skill between investors that may also be associated with bias exposure. We find that accounting for investor skill heterogeneity drops the coefficient by about one-third, though it remains economically and statistically significant at -1.28% with a t-statistic of -7.58. Time-invariant differences in skill between investors cannot explain the return differentials associated with the average bias. Instead, biased trading behavior, not the overall skill level of the individual trader or their level of experience, is likely driving the underperformance.

In columns 4 and 5, we assess whether the documented bias-return relationship differs from what we observe in the simulated data set.[6] This allows us to assess whether a bias toward certain stock characteristics is driving the underperformance. If we find a similar relationship in the simulated data, this would suggest that more biased trades underperform because of the types of stocks involved (i.e., loading up on particular stock characteristics). We use a single set of simulated data for this analysis (instead of the ten we used to assess relative prevalence).

In column 4, we only include the simulated trading data and find no relationship between bias exposure and future returns. The coefficient is 6 basis points and is insignificant. In

[6]In the main analyses, we use a single simulation to compare with the real data. In untabulated results, we confirm that our findings are robust to the choice of simulation.

column 5, we test whether the difference in the relationship between the real and simulated data is statistically significant. We include both real and simulated data. The coefficient of interest is on the interaction term between the indicator of real trading data and *Number of Biases*. We absorb the average return performance for each investor-data subset pair (Investor×Subset), so that each investor has a separate fixed effect for their trades in the real data and their trades in each simulated dataset. We also interact the control variables with a dummy indicator for the real dataset. We find that the coefficient is -0.0116 with a t-statistic greater than 6. The results in the last two columns suggest that the general bias-return relationship is not driven by bias toward particular stock characteristics.

In Figure 4, we plot the relationship between the number of biases and future returns from 1 month to 2 years (1 month, 3 months, 6 months, 9 months, 12 months and 24 months). We plot the coefficient on *Number of Biases* for two specifications: the specification in column 3 of Table 4, and the specification in column 5. We find that biases are not costly in the short-term with no impact up the 3 months. After 3 months, the relative returns turn negative and significant and continue to decline out to 2 years after the trade date. Considering the median holding date in our sample is about 26 months, biases are likely to negatively impact investor's actual realized return on their trades.

Overall, the results so far suggest that the average trading bias leads to future return underperformance in the medium- to long-run. This average underperformance is not explained by systematic risk of the stocks, the size of the trades, the investor's experience, the number of stocks held, or time-invariant differences in skill between investors. The underperformance is also not driven by bias towards particular stock characteristics associated with biases. Instead, the results suggest that the underperformance is likely due to poor stock selection or poor trade timing associated with bias exposure.

We next assess to what extent the timing of trades can explain the relative underperformance in Table 5. In column 1, we examine market-adjust returns instead of

raw returns. This will adjust for both the selected stock's market risk and the market return. We find that the coefficient on *Number of Biases* is unchanged compared to the coefficient in column 3 of Table 4. Market-adjusting returns does not impact the results suggesting market timing is not driving the return differential. We further explore this in column 2. We examine raw returns and add date fixed effects. The date fixed effects account for the timing of trades and absorb the average return of all trades made on that particular date. Accounting for the date of the trade leads us to estimate a slightly more negative impact of each additional bias on future returns.

In column 3, we include the simulated data, and continue to include the date fixed effects along with investor×subset fixed effects and controls. We find that the coefficient on the interaction term is -0.0097 with a t-statistic of -6.51. In column 4, we include Investor×Date fixed effects instead of date fixed effects, effectively comparing trades for the individual investor to their counterpart simulated trades in the simulated datasets. We find a similar coefficient of -0.0113 on the interaction term with a t-statistic of -6.58. The comparisons to the simulated trading data suggest that market timing and stock characteristics associated with biases combined only explain a small proportion of the negative relationship. Overall, the results suggest that the average bias associated with stock selection leads to a worse decision in terms of future performance.

5.1.1 Institutional Investor Biases and Future Performance

Do institutional investors perform similarly on bias-related trades? We examine the relationship between bias exposure and future returns for institutional investors in Table 6. In Panel A, we examine the relationship between bias exposure and performance for the eighteen biases that are commonly used by retail investors. We find that there is a negative relationship between biases and performance for institutions, thought it is not statistically significant in column 2 after including controls and investor fixed effects. In columns 3 and

4, we compare the simulated data and find that an additional bias in actual institutional investor trades is more costly than in the simulated data.

While the set of eighteen biases are generally costly to institutional investors, the institutions only exhibit six biases. In Panel B, we focus on the six biases exhibited by institutional investors. We find statistically weak evidence that this set of biases is costly to institutional investors.

In sum, biases in institutional investor trades are not as costly since institutions are able to avoid the costs by not suffering from as many biases.

In the remainder of the paper, we focus our analysis on retail investors to better understand the associated costs and their biased trading behavior.

5.2 Individual Biases and Future Returns

The previous results suggest that the average bias has a negative impact on investor performance. We next unpack this result by examining individual biases. In Figure 5, we present estimates of the future returns associated with each bias. For this analysis, we run a single regression with indicators for each individual bias, so the coefficient estimate for each bias is conditional on the other biases.

We find that seven out of eighteen individual biases exhibit a negative baseline relationship with future returns, five are not significantly related to future returns, and five are positively related to future returns. Four of the seven biases that show negative baseline performance are related to prior stock returns (*Max Return*, *Contrarian*, *Return Sorting Bottom*, *Return Sorting Top*). The other three are *Easy Name*, *Abnormal Trading Volume*, and *Peer Influence Industry*. Most of the biases with negative associated performance have been related to limited investor attention suggesting investors are purchasing stocks that grab their attention instead of rationally selecting stocks to add to their portfolio. Of the

five biases that show outperformance, two are related to transaction values (*Naive Buying Diversification* and *Integer Price*), the others are related to stock characteristics: *Volatility*, *Momentum*, and *High Dividend*.

After the inclusion of investor fixed effects and our base control variables, the estimated magnitudes of the coefficients for almost all biases decrease. All of the seven underperforming biases continue to exhibit underperformance. In addition, *Top Industry Experience* becomes underperforming. Four biases exhibit a positive and significant relationship once we account for time-invariant differences between investors and the base set of controls with *Naive Buying Diversification* no longer statistically significant.

We next adjust for timing and stock characteristics by estimating relative performance compared to simulated trades made on the same day that also satisfy the bias criteria. These tests will capture how investors fare on their biased stock purchases relative to randomly chosen stocks purchases made on the same day with similar characteristics such that the randomly chosen stock also satisfies the criteria of the bias. For example, the *High Dividend* estimate is the relative performance of investors' actual purchases of high-dividend stocks compared to the performance of a randomly selected stock that also has high-dividends purchased on the same day.

We find that in this specification there tends to be further attenuation of the estimated coefficients. For some biases, there is a reversal in sign. The general attenuation suggests that many of the biases exhibit under- or outperformance partially or entirely because of the timing of their trades or a bias toward particular characteristics. The biases that underperform in this set-up are: *High Dividend*, *Volatility*, *Top Industry Experience*, *Chase Dividend* , *Holding Period*, *Abnormal Trading Volume*, *Easy Name*, and *Max Return*. The biases that outperform are: *Contrarian*, *Integer Price*, and *Momentum*.

In general, the results suggest that while the average bias leads to poor performance, there is heterogeneity between biases in their impact on performance, with some leading to

outperformance and others leading to underperformance.

5.3 The Relationship Between Stock Selection Biases and Market Conditions

Our analysis thus far examines the average prevalence and average returns associated with our set of biases over the entire sample period. But investing is dynamic, investors learn by trading Seru et al. (2010), there are shifts in sentiment (Baker and Wurgler, 2007), and market conditions change, which may lead investors to adapt their stock selection methodologies. Given this, we next ask whether the prevalence and relative returns associated with our set of biases are related to market conditions. Our 15 year time series captures two market ups and downs in the Finnish market (Keppo et al., 2021), giving us ample variation in market conditions.

We begin by examining how the relationship between biases and returns differs across market conditions. In Table 7, we regress the 12 month returns of purchased stocks on *Number of Biases* and its interaction with market variables including the dividend-yield of the market index (*DY ratio*) , the past 12-month market return (*Past Year HEX Return*), and the past 12 month daily return volatility (*Past Year HEX Volatility*).

We find a significant relationship between market conditions and the costs associated with trading biases. In general, biases are less costly in "good" times, when valuations are higher, past market returns are higher, and past market volatility is lower. In "bad" times, biases are especially costly. For instance, if the past market volatility is one standard deviation greater, this increases the cost of an additional bias by 98.6% (from -0.0144 to -0.0286). These results suggest that investors biases do not cost them much in sanguine and rising markets, but are especially costly when markets are turbulent and declining. This time-variation in costs can delay learning about the true costs associated with a bias.

We next examine whether the prevalence of biases also changes across market conditions in Table 8. We estimate a Poisson regression with the number of biases associated with a trade as the dependent variable. We find that there is variation in the prevalence of biases over time. With prevalence increasing when the dividend-yield is higher, past returns have been lower or past return volatility is greater. These are precisely the times when the costs are greatest, suggesting there is a greater prevalence of biases when the exposure to them is the most costly.

Overall, these results highlight the ebbs and flows of biases in retail investor trading. Retail investors are more subject to trading biases in "bad" markets, and these biases exacerbate their underperformance. This pattern reflects either retail investors change their stock selection methodologies depending on the state of the market or the underlying biases of their stock selection methodologies change as the state of the market changes.

6 Understanding Biased Stock Selection

In this section, we discuss the costs and benefits of biased stock selection methods and empirically examine who are more likely to used biased methods for stock selection. We first provide a discussion and quantification of the performance costs documented in Section 5. We then provide a simple model of the factors affecting an individual's decision to continue to use a biased stock selection method. Lastly, we relate trading biases to investor demographics.

6.1 How Costly are Trading Biases?

The performance results suggest that the use of certain bias-related stock selection methodologies is especially costly for individuals in terms of lower future returns. To provide some additional color on the costs of biased selection methods, we perform a simple

back-of-the-envelope calculation using the coefficient of -0.0128 in column 3 of Table 4. Using the average trade size of 6,032.82 euros and a 12-month holding period, a one-standard deviation greater exposure to biases is associated with 77 euros of lower relative wealth ($-1.28\% \times 6,032.824$ euros) after 12 months. For comparison, the average monthly wage in Finland in Q1 2000 was 2,010 euros per month (Statistics Finland), suggesting that the cost is approximately six times the average worker's hourly wage.[7]

6.2 A Simple Theoretical Framework for Understanding Biased Trading Models

Given the average cost, why do retail investors use biased trading models to select stocks? A rational trader should trade off the potential costs (or benefits) in terms of return performance with potential benefits such as lower search costs and greater utility from the exposure to a biased model. We build a simple model that incorporates these features while allowing the investor to learn about the relative performance of the bias. Solving the model provides predictions on the conditions under which an investor will continue to use a biased trading method or switch to another method for asset selection.

The model has a single investor who buys a new single risky asset each period.[8] The investor is endowed with a stock selection method associated with a bias that she can use for asset selection. In each period, she decides to use either her endowed method to select an asset or use an alternative "benchmark" method, which can be thought of as an unbiased stock selection method such as purchasing the market portfolio. Each period, the investor decides to either keep or drop the endowed biased method. Once she decides to abandon the method, she will not be able to pick it up later. Our focus is on the drivers of the investor's stopping rule for discontinuing the exposure to a biased method.

[7] Source: https://tradingeconomics.com/finland/wages

[8] Our key observations from the model do not rely on this one-asset-per-period assumption, since the asset and period in the model can be generalized and interpreted as a portfolio and holding period.

For simplicity, we assume that the selected asset always follows some normal distribution fully characterized by its mean and variance parameters. The benchmark method leads to a selected asset whose return follows $\xi \sim \mathcal{N}\left(\mu_\xi, \sigma_\xi^2\right)$ with both μ_ξ and σ_ξ^2 observed by investors. The biased method leads to a selected asset whose return has a premium in excess of the benchmark method, that is, $r_h = \alpha + \xi$, where α is an unobserved constant. Since investors do not observe the value of α, they form subjective beliefs about it and update each period based on the return realized on their selected stock (a noisy signal). Formally, we use $\alpha \sim \mathcal{N}\left(m_{i,t}, s_{i,t}^2\right)$ to denote investor i's posterior belief about α in period t after observing the realized return in that period. Importantly, for all investors, the subjective distribution of α is independent of that of ξ. For simplicity, we assume that investors have an initial prior subjective belief in α with a zero mean.

We model the biased trading method as both costly and beneficial. The cost operates through a negative α, and the benefit is represented by a stream of utility $b_h > 0$, which can be either the savings in time and effort of using the biased method or the happiness obtained from the bias per se. We assume that investors have a mean variance preference for the next period return in addition to b_h from the method. Specifically, let $U_{i,t}$ denote investor i's myopic utility for the next period, which is given by:

$$U_{i,t} = \max\left\{U_{i,t}^{\text{drop}}, U_{i,t}^{\text{keep}}\right\}$$

s.t.

$$U_{i,t}^{\text{drop}} = \mathbb{E}\left[\xi\right] - 2\gamma \text{Var}\left[\xi\right] \tag{1}$$

$$U_{i,t}^{\text{keep}} = \widetilde{\mathbb{E}}_{i,t}\left[\xi + \alpha\right] - 2\gamma\widetilde{\text{Var}}_{i,t}\left[\xi + \alpha\right] + b_h, \tag{2}$$

where $U_{i,t}^{\text{drop}}$ is the utility from dropping the biased method in period t's stock selection and $U_{i,t}^{\text{keep}}$ is the utility from keeping the biased method in that period.

Comparing (1) and (2), we find that any investor i will drop the biased method if and only if

$$\begin{aligned} & U_{i,t}^{\text{drop}} > U_{i,t}^{\text{keep}} \\ \Leftrightarrow & -\underbrace{\widetilde{\mathbb{E}}_{i,t}[\alpha]}_{\equiv m_{i,t}} + 2\gamma \underbrace{\widetilde{\text{Var}}_{i,t}[\alpha]}_{\equiv s_{i,t}^2} > b_h, \end{aligned} \tag{3}$$

where $m_{i,t}$ and $s_{i,t}^2$ characterize i's posterior subjective belief about α. For the belief updating process, we let investors use standard Bayesian learning, which features

$$m_{i,t} = \frac{s_{i,t-1}^2}{s_{i,t-1}^2 + \sigma_\xi^2}(r_{i,t} - \mu_\xi) + \frac{\sigma_\xi^2}{s_{i,t-1}^2 + \sigma_\xi^2} m_{i,t-1} \tag{4}$$

$$\widetilde{s}_{0,t}^2 = \frac{s_{i,t-1}^2 \sigma_\xi^2}{s_{i,t-1}^2 + \sigma_\xi^2}. \tag{5}$$

where $r_{i,t}$ is the realized return, as the signal for learning, specific to investor i in period t from the stock selected by her in period $t-1$.

With the stopping rule in (3) and the learning rule in (4)-(5), we perform a simulation analysis to illustrate how biased method exposure is affected by the parameters of interest. Specifically, we focus on three parameters: (1) α: the true but unobserved return premium of the bias; (2) b_0: the benefit from using the bias; and (3) σ_ξ: the uncertainty of the benchmark asset selection method, which is the noise in the signal for learning. We consider three values for each of these parameters: $\alpha \in [-1.92\%, -1.28\%, -0.64\%]$, $\sigma_\xi \in [15.919\%, 31.837\%, 47.756\%]$, and $b_h \in [0.637\%, 1.274\%, 1.911\%]$, where the midpoint for σ_ξ is calibrated to the standard deviation of the return of the HEX index, the midpoint for α is obtained from column 3 in Table 4, and the midpoint for b_h is calibrated to one-tenth of the average return of the HEX index. The other two values for each parameter are obtained by adding or subtracting half of the midpoint value. For each combination of these parameter values, we run 100 independent simulations with each having 1,000 investors for 100 periods.

Figure 6 shows that a higher level of utility (b_h) leads to a more persistent exposure to the endowed biased method, which is expected given how b_h appears in the stopping rule in (3). In addition, a lower (or more negative) bias premium (α) leads to a slightly faster drop rate (that is, lower persistence). The effect on persistence is less pronounced than that of the utility level, although it has a magnitude similar to that of b_h. This is because the bias premium (α) is not directly observed and can only play a role in investors' learning and decision making indirectly through the noisy realized return as a signal. Within each plot, a higher level of uncertainty in the benchmark return (σ_ξ), i.e., a greater noise in the signal, corresponds to a greater persistence in biased method exposure. This is intuitive since noisier signals make it harder for investors to detect and attribute the negative premium to the exposure to biases.

The model formalizes the intuition that investors will take time to learn about the costs associated with their (biased) model. To see why, the bias premium α is relatively higher, or even positive, during good times, making it less likely that investors will drop the biased method based on our model and simulation results during good times. During bad times, on the other hand, our model and simulation suggest that higher market noise makes it more difficult to attribute losses to biases, even though the bias premium α is significantly negative.

While it may be difficult for investors to quantify the costs, even if they understood the costs it is not clear whether using a biased strategy is suboptimal. With optimality also depending on the associated benefits of using these methods. We cannot rule out the possibility that the investors enjoy lower information processing costs by using these biased selection methods. The above calculation gives some guidance on how large these costs must be to justify their continued exposure.

Although investors likely derive some utility from using certain biased methods (for example, the exposure to the *Max return* bias may relate to preferences for lottery-type

stocks), it is not clear how the magnitude of this utility compares to that of the financial loss from lower future returns. Moreover, for most biases in our analysis, we cannot find a reasonable argument that they contribute to a higher utility level.

6.3 Who Uses More Biased Stock Selection?

Finally, we examine how biased stock selection correlates with investor demographic variables. We perform this analysis at the trade level. The demographics of the investors we examine are gender, age and years of experience. We estimate a Poisson regression with the number of biases associated with a trade as the dependent variable. The explanatory variables of interest are the demographic characteristics and their interaction with market condition variables, including dividend-yield of the market index, the past 12-month market return, and the past 12 month daily return volatility.

The results are presented in Table 9. The results in column (1)-(3) suggest that biases are not used uniformly across investors and that, surprisingly, they are generally used more often by more sophisticated investors with greater years of experience. We also find differences in bias exposure related to gender. Females are more likely to trade with more biases, with 1.34% ($= e^{0.0134} - 1$) more biases per trade on average.

We also consider the interactions of characteristic variables with certain market conditions. In column (4), we show that when dividend-yield is high (i.e., current market index level is low), the effect of years of experience on bias exposure is weakened. This is substantiated by the result in column (5) as a low past market return leads to a weakened years of experience effect. These results suggest that more experienced investors make less biased tradings during bad times compared to what they would do during good times. This observation is further verified in column (6), which shows that more experienced investors' tradings display less biases under more volatile market conditions.

Overall, these results suggest that although more sophisticated investors are also biased

in their trading, they reduce bias exposure when the market condition worsens.

7 Conclusion

We study the performance of a broad set of trading biases using 14.5 years of retail and institutional investor trading data. Our set of biases is derived from processing articles published in top finance journals over the past 75 years. From an initial set of twenty-two biases, we find that around 20% of the identified biases are not used more frequently by retail investors than would be expected in randomized trading data.

We find a negative relationship between trading biases and future returns in retail investors' trades. Purchases involving these biases systematically underperform over the following year, with heterogeneity in the association between individual biases and future returns. This underperformance persists for up to two years. In contrast, institutions use only six of the twenty-two biases more frequently than would be expected under random trading and mitigate underperformance by selectively employing this limited subset.

We find that the prevalence of biases among retail investors varies with market conditions. Their trades exhibit more biases in "bad times", when the associated underperformance is especially severe.

To better understand retail investors' trading biases, we develop a model and conduct cross-sectional analyses. The model formalizes the intuition that investors find it difficult to recognize the costs of biases. We also examine how biased stock selection correlates with demographics, finding that while sophisticated investors also exhibit biases, they reduce their exposure when market conditions worsen.

Overall, our study provides systematic evidence on the prevalence, the costs, and conditional nature of stock selection biases in retail and institutional investors. The findings advance our understanding of how trading biases shape investment outcomes and underscore

the importance of considering market conditions when assessing their costs.

References

Baker, Malcolm, and Jeffrey Wurgler, 2007, Investor sentiment in the stock market, *Journal of economic perspectives* 21, 129–151.

Bali, Turan G, Nusret Cakici, and Robert F Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of financial economics* 99, 427–446.

Barber, Brad M, Xing Huang, Terrance Odean, and Christopher Schwarz, 2022, Attention-induced trading and returns: Evidence from robinhood users, *The Journal of Finance* 77, 3141–3190.

Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *The Journal of Finance* 55, 773–806.

Barber, Brad M., and Terrance Odean, 2001, Boys will be boys: Gender, overconfidence, and common stock investment*, *The Quarterly Journal of Economics* 116, 261–292.

Barber, Brad M, and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *The review of financial studies* 21, 785–818.

Barber, Brad M, Terrance Odean, and Ning Zhu, 2008, Do retail trades move markets?, *The Review of Financial Studies* 22, 151–186.

Birru, Justin, and Baolian Wang, 2016, Nominal price illusion, *Journal of Financial Economics* 119, 578–598.

Brunnermeier, Markus, and Stefan Nagel, 2004, Hedge funds and the technology bubble, *The Journal of Finance* 59, 2013–2040.

Cosemans, Mathijs, and Rik Frehen, 2021, Salience theory and stock prices: Empirical evidence, *Journal of Financial Economics* 140, 460–483.

Coval, Joshua D, David Hirshleifer, and Tyler Shumway, 2021, Can individual investors beat the market?, *The Review of Asset Pricing Studies* 11, 552–579.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In search of attention, *The journal of finance* 66, 1461–1499.

Da, Zhi, Xing Huang, and Lawrence J Jin, 2021, Extrapolative beliefs in the cross-section: What can we learn from the crowds?, *Journal of Financial Economics* 140, 175–196.

Daniel, Kent, Lorenzo Garlappi, and Kairong Xiao, 2021, Monetary policy and reaching for income, *The Journal of Finance* 76, 1145–1193.

Davis, Carter, Samuli Knüpfer, Jens Soerlie Kvaerner, Bahar Sen Dogan, and Petra Vokata, 2025, Do households matter for asset prices?, *Fisher College of Business Working Paper* .

Dorn, Daniel, and Gur Huberman, 2010, Preferred risk habitat of individual investors, *Journal of Financial Economics* 97, 155–173.

Eaton, Gregory W, T Clifton Green, Brian S Roseman, and Yanbin Wu, 2022, Retail trader sophistication and stock market quality: Evidence from brokerage outages, *Journal of Financial Economics* 146, 502–528.

Eraker, Bjørn, and Mark Ready, 2015, Do investors overpay for stocks with lottery-like payoffs? an examination of the returns of otc stocks, *Journal of Financial Economics* 115, 486–504.

Escobar, Laura, and Alvaro Pedraza, 2023, Active trading and (poor) performance: The social transmission channel, *Journal of Financial Economics* 150, 139–165.

Feng, Lei, and Mark S Seasholes, 2004, Correlated trading and location, *The Journal of finance* 59, 2117–2144.

Gargano, Antonio, and Alberto G Rossi, 2018, Does it pay to pay attention?, *The Review of Financial Studies* 31, 4595–4649.

Gathergood, John, David Hirshleifer, David Leake, Hiroaki Sakaguchi, and Neil Stewart, 2023, Naive buying diversification and narrow framing by individual investors, *The Journal of Finance* 78, 1705–1741.

Gigerenzer, Gerd, 2008, *Rationality for mortals: Risk and rules of thumb* (Oxford University Press, New York, NY).

Gigerenzer, Gerd, and Henry Brighton, 2009, Homo Heuristicus: Why Biased Minds Make Better Inferences, *Topics in Cognitive Science* 1, 107–143.

Graham, John R, and Alok Kumar, 2006, Do dividend clienteles exist? evidence on dividend preferences of retail investors, *The Journal of Finance* 61, 1305–1336.

Green, T Clifton, and Byoung-Hyoun Hwang, 2009, Price-based return comovement, *Journal of financial economics* 93, 37–50.

Green, T Clifton, and Russell Jame, 2013, Company name fluency, investor recognition, and firm value, *Journal of Financial Economics* 109, 813–834.

Grinblatt, Mark, and Matti Keloharju, 2000, The investment behavior and performance of various investor types: a study of finland's unique data set, *Journal of financial economics* 55, 43–67.

Grinblatt, Mark, and Matti Keloharju, 2001a, How distance, language, and culture influence stockholdings and trades, *The Journal of Finance* 56, 1053–1073.

Grinblatt, Mark, and Matti Keloharju, 2001b, What makes investors trade?, *The journal of Finance* 56, 589–616.

Grinblatt, MARK, and MATTI Keloharju, 2009, Sensation seeking, overconfidence, and trading activity, *The Journal of Finance* 64, 549–578.

Grinblatt, Mark, Matti Keloharju, and Juhani T. Linnainmaa, 2012, Iq, trading behavior, and performance, *Journal of Financial Economics* 104, 339–362, Special Issue on Investor Sentiment.

Hartzmark, Samuel M, 2015, The worst, the best, ignoring all the rest: The rank effect and trading behavior, *The Review of Financial Studies* 28, 1024–1059.

Hartzmark, Samuel M, and David H Solomon, 2019, The dividend disconnect, *The Journal of Finance* 74, 2153–2199.

Huang, Shiyang, Byoung-Hyoun Hwang, and Dong Lou, 2021, The rate of communication, *Journal of Financial Economics* 141, 533–550.

Huang, Xing, 2019, Mark twain's cat: Investment experience, categorical thinking, and stock selection, *Journal of Financial Economics* 131, 404–432.

Huberman, Gur, 2001, Familiarity breeds investment, *The Review of Financial Studies* 14, 659–680.

Hvide, Hans K, and Per Östberg, 2015, Social interaction at work, *Journal of Financial Economics* 117, 628–652.

Ivković, Zoran, and Scott Weisbenner, 2005, Local does as local is: Information content of the geography of individual investors' common stock investments, *The Journal of Finance* 60, 267–306.

Ivković, Zoran, and Scott Weisbenner, 2007, Information diffusion effects in individual investors' common stock purchases: Covet thy neighbors' investment choices, *The Review of Financial Studies* 20, 1327–1357.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *The Journal of finance* 48, 65–91.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, 2008, Individual investor trading and stock returns, *The Journal of finance* 63, 273–310.

Kaustia, Markku, and Samuli Knüpfer, 2008, Do investors overweight personal experience? evidence from ipo subscriptions, *The Journal of Finance* 63, 2679–2702.

Keloharju, Matti, Samuli Knüpfer, and Juhani Linnainmaa, 2012, Do investors buy what they know? product market choices and investment decisions, *The Review of Financial Studies* 25, 2921–2958.

Keppo, Jussi, Tyler Shumway, and Daniel Weagley, 2021, Are monthly market returns predictable?, *The Review of Asset Pricing Studies* 11, 806–836.

Korniotis, George M., and Alok Kumar, 2011, Do older investors make better investment decisions?, *The Review of Economics and Statistics* 93, 244–265.

Kumar, ALOK, 2009, Who gambles in the stock market?, *The Journal of Finance* 64, 1889–1933.

Kumar, Alok, and Charles MC Lee, 2006, Retail investor sentiment and return comovements, *The Journal of Finance* 61, 2451–2486.

Kuo, Wei-Yu, Tse-Chun Lin, and Jing Zhao, 2015, Cognitive limitation and investment performance: Evidence from limit order clustering, *The Review of Financial Studies* 28, 838–875.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny, 1994, Contrarian investment, extrapolation, and risk, *The journal of finance* 49, 1541–1578.

Liao, Jingchi, Cameron Peng, and Ning Zhu, 2022, Extrapolative bubbles and trading volume, *The Review of Financial Studies* 35, 1682–1722.

Liu, Bibo, Huijun Wang, Jianfeng Yu, and Shen Zhao, 2020, Time-varying demand for lottery: Speculation ahead of earnings announcements, *Journal of Financial Economics* 138, 789–817.

Malmendier, Ulrike, 2021, Experience effects in finance: Foundations, applications, and future directions, *Review of Finance* 25, 1339–1363.

Massa, Massimo, and Andrei Simonov, 2006, Hedging, familiarity and portfolio choice, *The Review of Financial Studies* 19, 633–685.

Neth, Hansjörg, and Gerd Gigerenzer, 2015, Heuristics: Tools for an uncertain world, in *Emerging trends in the social and behavioral sciences*, 1–18 (Wiley Online Library).

Pool, Veronika K, Noah Stoffman, and Scott E Yonker, 2015, The people in your neighborhood: Social interactions and mutual fund portfolios, *The Journal of Finance* 70, 2679–2732.

Rantapuska, Elias, 2008, Ex-dividend day trading: Who, how, and why?: Evidence from the finnish market, *Journal of Financial Economics* 88, 355–374.

Seasholes, Mark S, and Ning Zhu, 2010, Individual investors and local bias, *The Journal of Finance* 65, 1987–2010.

Seru, Amit, Tyler Shumway, and Noah Stoffman, 2010, Learning by trading, *The Review of Financial Studies* 23, 705–739.

Tversky, Amos, and Daniel Kahneman, 1974, Judgment under uncertainty: Heuristics and biases, *Science* 185, 1124–1131.

Welch, Ivo, 2022, The wisdom of the robinhood crowd, *The Journal of Finance* 77, 1489–1527.

Figure 1: **Time-Series of Published Papers on Biases**

The figure shows the number of papers published in the top finance journals each year. We present the total number of papers, the number of papers classified as behavioral finance, and the number of papers on trading biases. The figure includes papers published in the *Journal of Finance*, *Journal of Financial Economics*, and the *Review of Financial Studies* from 1949 to 2023.



Figure 2: **Histogram of the Number of Biases Used Per Trade**

The figure compares the density distributions of *Number of Biases*, which is the total count of biases associated with a single trade We plot real buy trades (blue bars), and simulated buy trades (orange hatched bars) generated using a simulation algorithm, based on a randomly selected sample of 50,000 investors. We focus only on the 18 biases that are more prevalent than what we find in the simulated data, as documented in Figure 3. In the real trading data, 90.51% of trades are associated with more than one bias, with a median of three biases per trade.

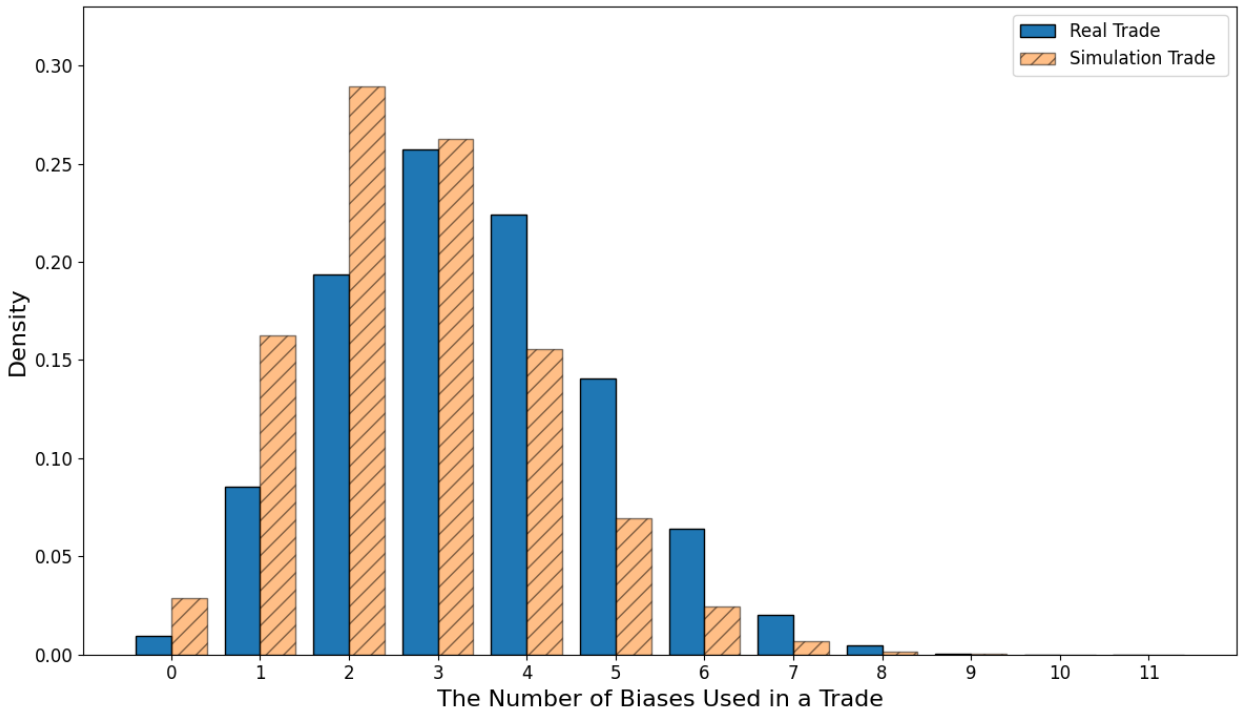


Figure 3: **The Prevalence of Biases in Real Trading Relative to Simulated Trading**

The figure compares the prevalence of each bias between real trading data and simulations generated by a simulation algorithm. *Prevalence* is measured by the ratio of trades associated with each bias to all trades in a sample. *Prevalence* is plotted in decimals along the x-axis. In Panel A, we use the entire trading records of a randomly selected sample of 50,000 retail investors who have at least two buy or sell transactions during the sample period. In Panel B, we use the entire trading records of 1,065 institutional investors over the same period. Circle markers show prevalence from real trading data, whereas cross markers show the 95% confidence interval of simulated prevalence (10 simulations of the algorithm). The 22 biases are ordered by the difference between real prevalence and the maximum of the simulation values. The light gray area represents biases that are less prevalent than in at least one version of the simulations.

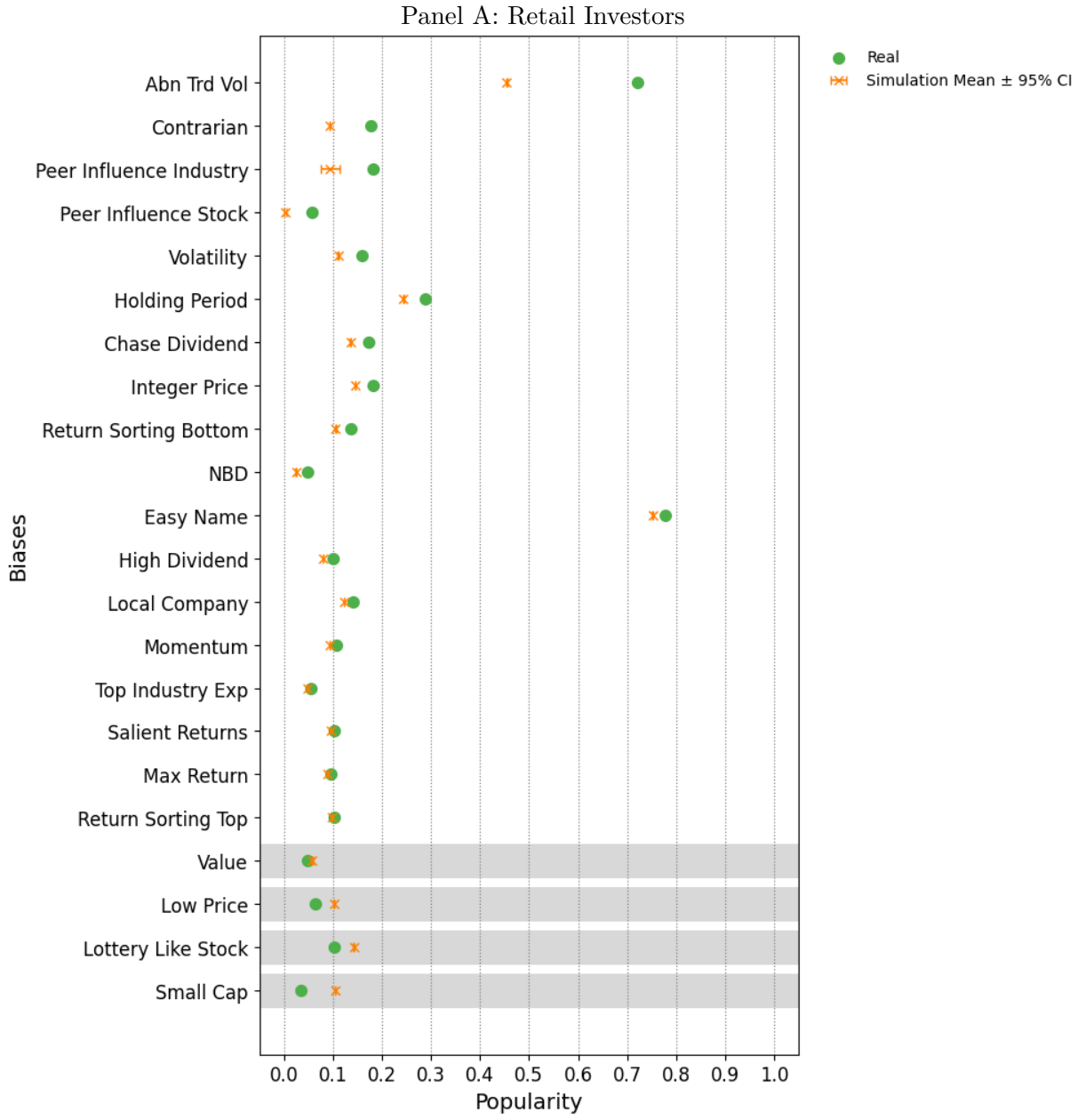


Figure 3 continued...

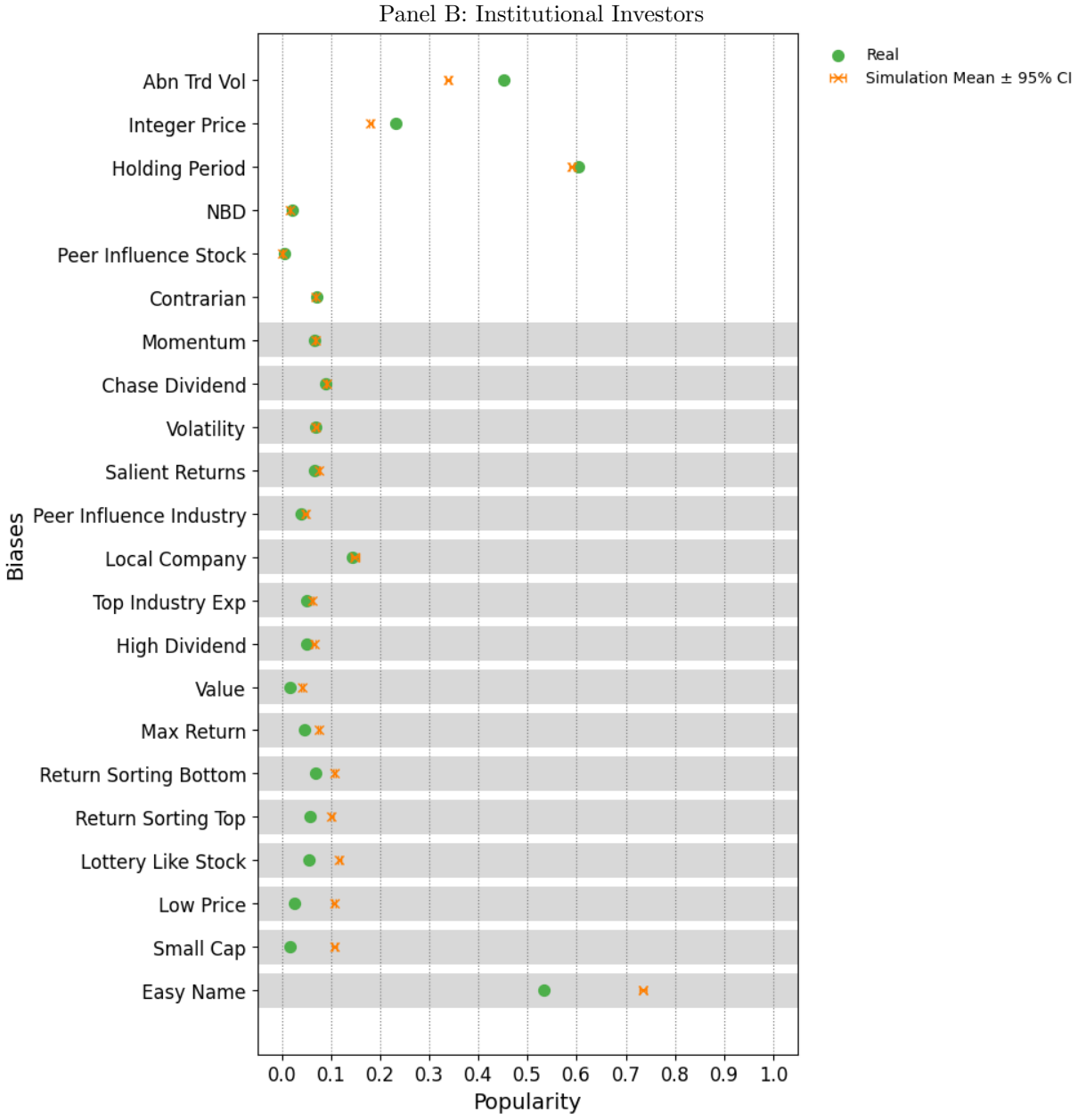


Figure 4: **The Performance of Biased Stock Selection over Time**

The figure illustrates the relationship between biases exposure and the future returns of the retail investors' trade. The dependent variable, *Future Return*, is the raw return of the stock traded over a specific time windows following the trade. In this figure we measure *Future Return* over 1, 3, 6, 9, 12, and 24-month windows. The main independent variable, *Number of Biases*, is defined as the total count of biases associated with a single trade. In this figure, we include 18 biases commonly used by retail investors. Circle markers show the coefficient estimates obtained from regressions of *Future Return* on *Number of Biases* without any controls and any fixed effects, using real trading data only. Square markers show the coefficient estimates for interaction terms between *Number of Biases* and *Real*, 1 if the trade is from the real trading sample and 0 if from the simulations, from regressions incorporating control variables, investor fixed effects, and investor-day fixed effects, and using both real and simulated data. Vertical bars denote 95% confidence intervals where standard errors are clustered at investor and day level. Control variables include *Stock Market Beta*, the logarithm of *Total Stocks*, the logarithm of *Trade Size*, *Years of Experience*, and the logarithm of *Age*.

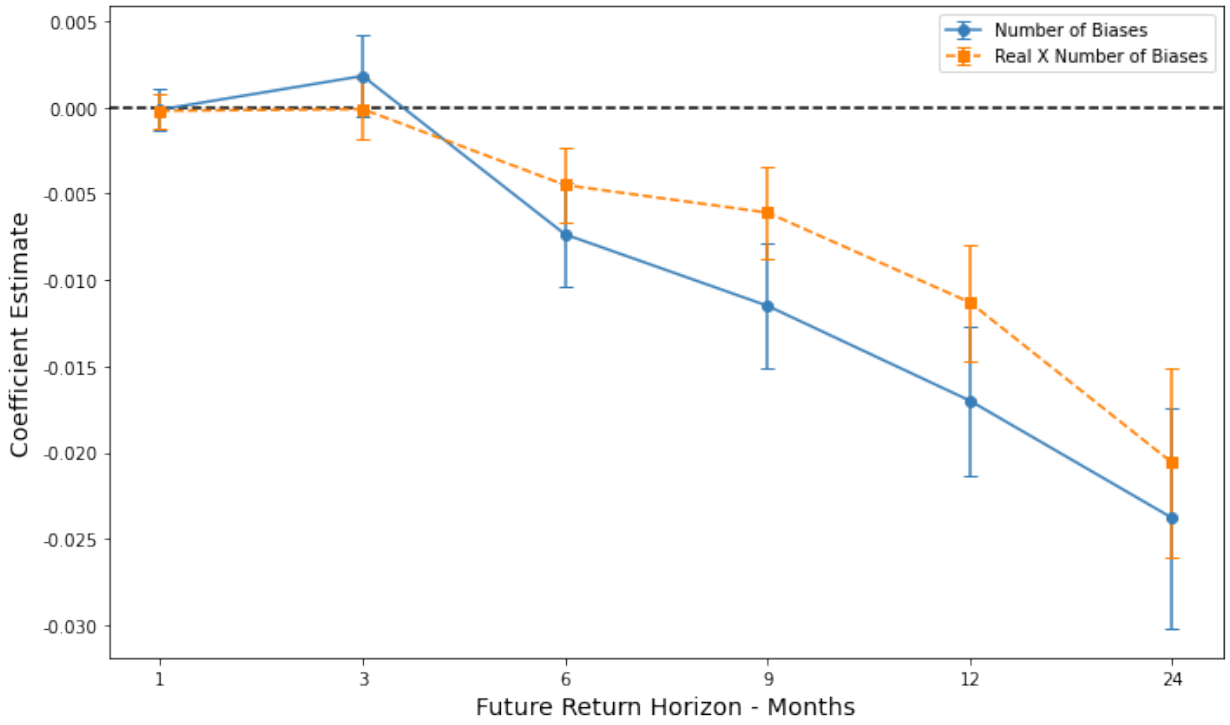


Figure 5: **The Relationship Between Individual Bias and Future Return**

The figure shows the relationship between individual bias exposure and trading performance. The dependent variable, *Future Return*, represents the raw return of the stock traded over a specific time windows following the trade, measured over a 12-month window following the trade. The main independent variable, $Bias_b$, is defined as an indicator for each trade, which equals one if the trade meets the definition of bias b. In this figure, we include 18 biases commonly used by retail investors. Circle markers represent coefficient estimates obtained from a baseline regression of *Future Return* on $Bias_b$ without any controls, using real trading data exclusively. Square markers denote coefficient estimates from a regression of *Future Return* on $Bias_b$, including control variables and investor fixed effects, also using real trading data. Diamond markers indicate coefficient estimates for interaction terms between $Bias_b$ and *Real*, defined as 1 if the trade is from the real trading sample and 0 if from the simulations, from regressions incorporating investor-day fixed effects and using both real and simulated data. All markers are shown with their 95% confidence intervals where standard errors are clustered at the investor and date levels. Control variables include *Stock Market Beta*, the logarithm of *Total Stocks*, the logarithm of *Trade Size*, *Years of Experience*, and the logarithm of *Age*.

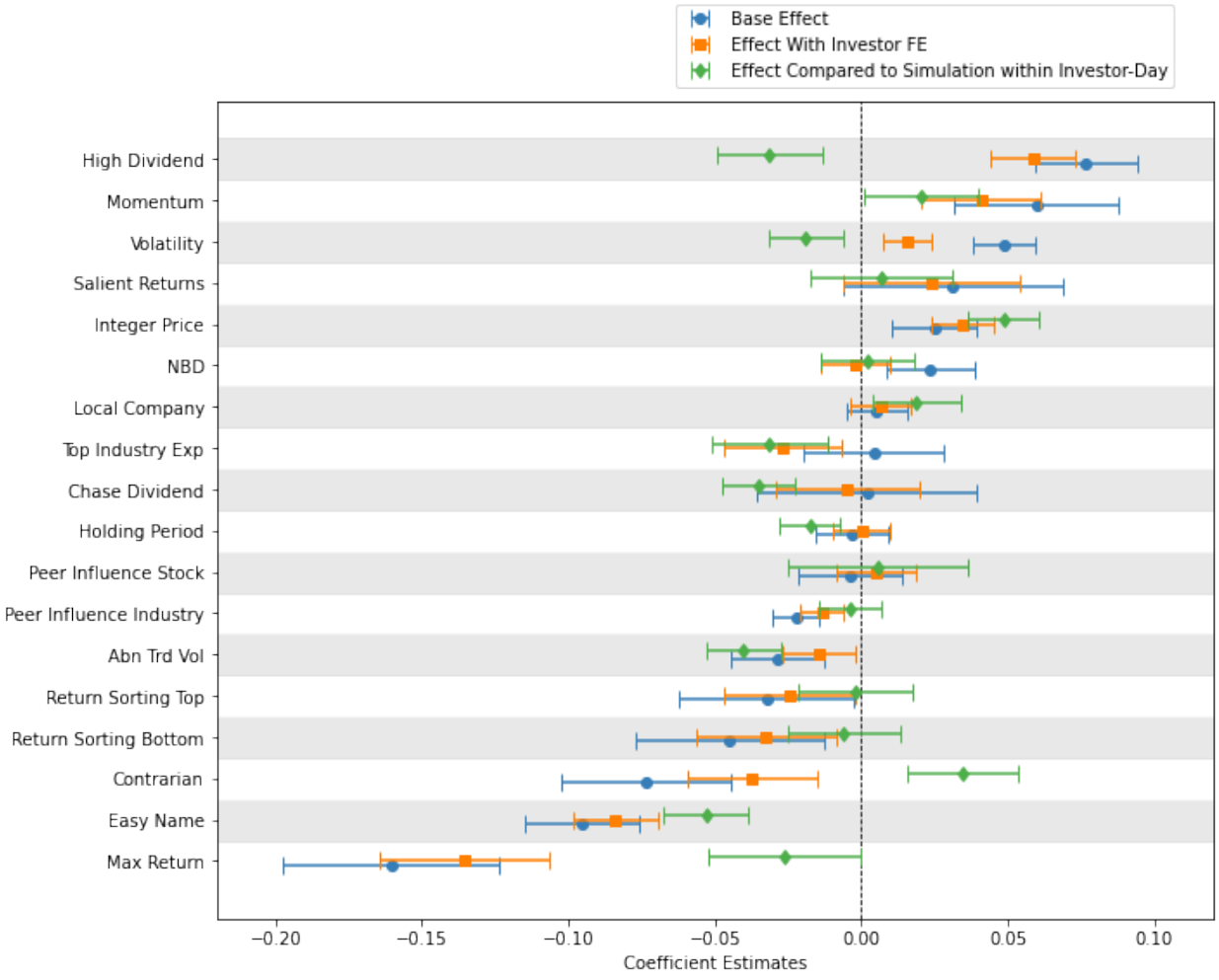


Figure 6: **Model Simulations**

The figure shows the fraction of investors who continue to have exposure to the endowed bias over time. The three columns, from left to right, correspond to low utility ($b_h = 0.637\%$), medium utility ($b_h = 1.274\%$), and high utility ($b_h = 1.911\%$); the three rows, from top to bottom, correspond to low bias premium ($\alpha = -1.680\%$), medium bias premium ($\alpha = -1.120\%$), and high bias premium ($\alpha = -0.560\%$). Within each plot, the blue solid line, red dashed line, and green dotted line correspond to low benchmark return standard deviation ($\sigma_\xi = 15.919\%$), medium benchmark return standard deviation ($\sigma_\xi = 31.837\%$), and high benchmark return standard deviation ($\sigma_\xi = 47.756\%$).

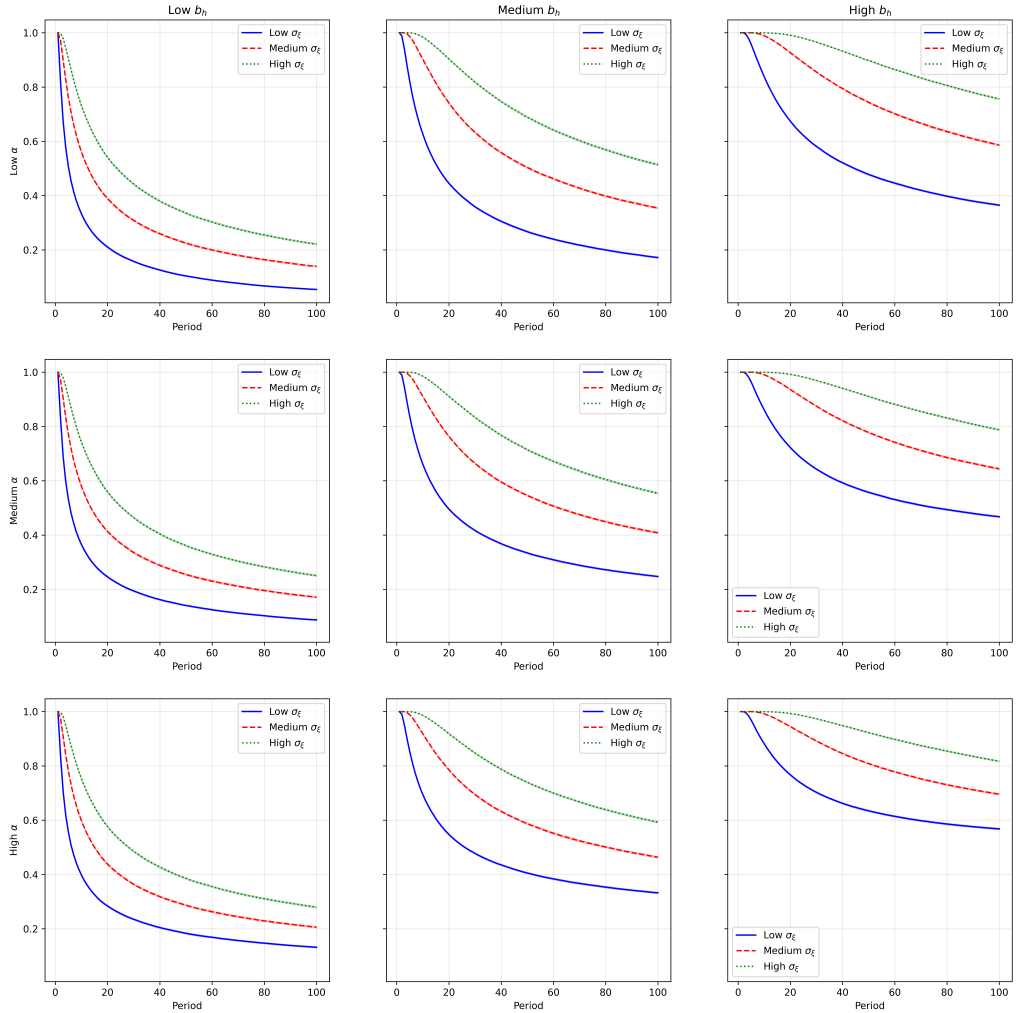


Table 1: **Paper Counts and Classification Results**

The table shows the number of papers investigating trading biases published in three finance journals: *Journal of Finance*, *Journal of Financial Economics*, and *Review of Financial Studies*. For each journal, we examine papers published during the specified time range and report the total number of papers in the column "All Papers". We then adopt the two-step determination procedure described in Subsection 2.2 and report the number of behavioral finance papers and the number of trading bias papers in the columns "Behavioral Papers" and "Trading Bias Papers", respectively. Our classification approach emphasizes minimizing Type II errors, ensuring the inclusion of nearly all relevant papers, which inevitably leads to many "false positive" results, meaning that some of these papers are either not about or only remotely related to individual investors and stock investment. We adopt a manual review process later to screen these 646 papers.

Journal	Time Range	All Papers	Behavioral Papers	Trading Bias Papers
The Journal of Finance	1946 - 2023	5105	920	208
Journal of Financial Economics	1974 - 2023	3682	1199	238
Review of Financial Studies	1988 - 2023	2517	729	200
Total		11,304	2,848	646

Table 2: **Result for 100-round Bootstrap**

This table shows the reliability of our paper classification process. We use a bootstrap methodology to evaluate the Type II error of paper classification. Across 100 iterations, we randomly select 10 papers in each round and present the same questions to the ChatGPT-4 model as in our main classification process.

	Average Accuracy	Average Type II error	Std for Type II error	95% CI for Type II error
Trading Bias	83.00%	17.00%	1.41%	[14.17%, 19.83%]
Behavioral Finance	100.00%	0.00%	0.00%	[0.00%, 0.00%]

Table 3: **Summary Statistics**

The table presents summary statistics for variables at the trade level. We include trade observations from the randomly selected sample of 50,000 investors. *Number of Biases* is the total count of biases associated with a single trade. For retail investors, we focus on 18 prevalent biases, as shown in Panel A of Table 3, and for institutional investors, we focus on 6 prevalent biases in Panel B. *Future Return* is the raw return of the stock traded over a specific time windows following the trade. *Future market-adjusted return* is the market-adjusted return of the stock traded over a specific time windows following the trade, i.e., stock excess return adjusted to stock market beta and excess return of HEX index. We mainly focus on 12-month returns. *Daily Market Return* is a daily HEX return in decimal. *Past Year Market Volatility* is Market volatility over the past 12 months, calculated using daily HEX prices in decimal. *Past Month Market Return* is the HEX return over the 21 trading days preceding the trade date in decimal. *Stock Market Beta* is the beta coefficient estimated by regressing the stock's excess weekly returns on the value-weighted excess market returns over the past 52 weeks.. *Total Stocks* is the number of stocks in an investor's portfolio. *Trade Size* is the principal value of a trade, expressed in euros and adjusted for inflation to constant euros of the year 2000. *Years of Experience* is the number of years since the investor's first trade to the current trade. *Trading Intensity* is calculated as the total number of buy transactions made before the current trade divided by the number of months the investor has been active in the market. *Female* is a binary indicator equal to 1 if the trade is made by a female trader and 0 if made by a male trader. *Age* is the trader's current age at the time of each trade.

Variable	Obs.	Mean	Std.Dev.	p25	p50	p75
Number of Biases (retail)	363,084	3.59	1.60	2.00	3.00	5.00
Number of Biases (institution)	343,111	1.42	0.97	1.00	1.00	2.00
12-month Returns						
Future Return	338,655	0.12	0.61	-0.28	0.06	0.38
Future market-adjusted Return	312,794	-0.02	0.53	-0.27	-0.06	0.21
Stock Market Beta	316,326	0.64	0.48	0.30	0.64	0.94
Total Stocks	363,084	7.72	8.27	3.00	6.00	10.00
Trade Size	363,084	6032.82	48843.79	928.30	2269.70	5344.14
Years of Experience	363,084	5.48	4.23	1.58	5.00	8.83
Female	362,830	0.20	0.40	0.00	0.00	0.00
Age	362,415	46.99	15.23	36.00	48.00	58.00
DY ratio	363,049	2.88	1.75	1.28	2.82	3.48
Past Year HEX Return	363,084	0.11	0.43	-0.24	0.16	0.31
Future Year HEX Return	363,084	0.06	0.36	-0.24	0.14	0.28
Past Year HEX Volatility	361,231	0.30	0.12	0.18	0.32	0.37
Future Year HEX Volatility	363,084	0.32	0.13	0.20	0.31	0.40

Table 4: **The Relationship Between Stock Selection Biases and Performance**

The table shows the relationship between the future return of a trade and bias exposure in the trade. The dependent variable, *Future Return*, is the raw return of the stock traded over a specific time windows following the trade. We use 12-month window after the trade. The main independent variable, *Number of Biases*, is the total count of biases associated with a single trade. From Columns 1 to 3, we focus on the real trading data. In Column 4, we focus on the simulated trading data. We use simulated trading data in Columns 4. In Column 5, we include the interaction terms of *Number of Biases* with *Real*, which is 1 if the trade is from the real trading sample and 0 if from the simulations. In all columns we control for *Stock Market Beta*, which is the beta coefficient estimated by regressing the stock's excess weekly returns on the value-weighted excess market returns over the past 52 weeks.. We also include the logarithm of *Total Stocks*, the logarithm of *Trade Size*, *Years of Experience*, and the logarithm of *Age* as control variables. Standard errors are clustered at the investor and date levels. t-statistics are in parentheses. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	(1)	(2)	(3)	(4)	(5)
Number of Biases	-0.0170***	-0.0194***	-0.0128***	-0.0013	-0.0013
	(-7.71)	(-9.15)	(-7.58)	(-1.03)	(-1.03)
Real × Number of Biases					-0.0116***
					(-6.65)
Data Subset	Real	Real	Real	Simulation	Both
Controls	No	Yes	Yes	Yes	Yes
Investor FEs	No	No	Yes	No	No
Investor×Subset FEs	No	No	No	Yes	Yes
Date FEs	No	No	No	No	No
Adj. R^2	0.0019	0.0139	0.1495	0.1083	0.1327
Obs.	338,655	315,957	309,839	285,680	595,519

Table 5: **The Relationship Between Stock Selection Biases and Performance: Adjusting for Market Timing**

The table shows the relationship between the future return of a trade and bias exposure in the trade. In Column 1, the dependent variable is *Future market-adjusted return* which is the market-adjusted return of the stock traded over a specific time windows following the trade, i.e., stock excess return adjusted to stock market beta and excess return of HEX index. In Column 2 to 4, the dependent variable is *Future Return*, the raw return of the stock traded over a specific time windows following the trade. We use a 12-month window after the trade. The main independent variable, *Number of Biases*, is the total count of biases associated with a single trade. In Column 1 and 2, we focus on the real trading data, while we combine the real and simulated trading data in Columns 3 and 4. In Columns 3 and 4, we include the interaction terms of *Number of Biases* with *Real*, which is 1 if the trade is from the real trading sample and 0 if from the simulations. In all columns, we control for *Stock Market Beta*, which is the beta coefficient estimated by regressing the stock's excess weekly returns on the value-weighted excess market returns over the past 52 weeks.. We also include the log of *Total Stocks*, the log of *Trade Size*, *Years of Experience*, and the log of *Age* as control variables. s Standard errors are clustered at the investor and date levels. t-statistics are in parentheses. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	(1)	(2)	(3)	(4)
Dependent Variable	Future market-adjusted Return		Future Return	
Number of Biases	-0.0108***	-0.0132***	-0.0036***	-0.0022*
	(-7.38)	(-9.02)	(-3.76)	(-1.79)
Real × Number of Biases			-0.0097***	-0.0113***
			(-6.51)	(-6.58)
Data Subset	Real	Real	Both	Both
Controls	Yes	Yes	Yes	Yes
Investor×Subset FEs	Yes	Yes	Yes	Yes
Date FEs	No	Yes	Yes	No
Investor×Date FEs	No	No	No	Yes
Adj. R^2	0.1319	0.4458	0.3862	0.3224
Obs.	306,624	309,839	595,519	510,777

Table 6: **The Relationship Between Stock Selection Biases and Performance: Institutional Investors**

The table shows the relationship between the future return of a trade and bias exposure in the trade using a sample of institutional investors. The dependent variable, *Future Return*, is the raw return of the stock traded over a specific time windows following the trade. We use a 12-month window after the trade. The main independent variable, *Number of Biases*, is the total count of biases associated with a single trade. In Panel A, we use a set of 18 biases prevalently used by retail investors, as shown in Figure 3, Panel A. In Panel B, we use a set of 6 biases prevalently used by institutional investors, as shown in Figure 3, Panel B. In Columns 3 and 4 of both panels, we include the interaction terms of *Number of Biases* with *Real*, which is 1 if the trade is from the real trading sample and 0 if from the simulations. When including control variables, we control for *Stock Market Beta*, which is the beta coefficient estimated by regressing the stock's excess weekly returns on the value-weighted excess market returns over the past 52 weeks.. We also include the log of *Total Stocks*, the log of *Trade Size*, and *Years of Experience*. Standard errors are clustered at the investor and date levels. t-statistics are in parentheses. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Panel A: 18 Retail Prevalent Biases

	(1)	(2)	(3)	(4)
Number of Biases	-0.0150**	-0.0073	0.0195*	0.0098**
	(-2.29)	(-1.52)	(1.74)	(2.09)
Real × Number of Biases			-0.0269***	-0.0097**
			(-3.07)	(-2.26)
Data Subset	Real	Real	Both	Both
Controls	No	Yes	Yes	Yes
Investor FEs	No	Yes	No	No
Investor×Subset FEs	No	No	Yes	Yes
Date FEs	No	No	No	Yes
Adj. R^2	0.0013	0.0689	0.0663	0.3705
Obs.	232,198	202,710	437,086	389,528

Panel B: 6 Institution Prevalent Biases

	(1)	(2)	(3)	(4)
Number of Biases	-0.0124	0.0086*	0.0011	-0.0048
	(-1.23)	(1.65)	(0.31)	(-1.27)
Real × Number of Biases			0.0075	0.0047
			(1.43)	(1.18)
Data Subset	Real	Real	Both	Both
Controls	No	Yes	Yes	Yes
Investor FEs	No	Yes	No	No
Investor×Subset FEs	No	No	Yes	Yes
Date FEs	No	No	No	Yes
Adj. R^2	0.0003	0.0687	0.0649	0.3701
Obs.	232,198	202,710	437,086	389,528

Table 7: **How the Performance of Biased Stock Selection Varies Across Market Conditions**

The table examines the relationship between bias exposure and the market conditions. The dependent variable, *Future Return*, is the raw return of the stock traded over a specific time windows following the trade. We use a 12-month window after the trade. The main independent variable is *Number of Biases* which is the total count of biases associated with a single trade. We include the interaction term between *Number of Biases* and market condition variables. We use several market condition variables. From Columns 1 to 5, we use *DY ratio*, *Past Year HEX Return*, and *Past Year HEX Volatility*, Future *Future Year HEX Return*, and *Future Year HEX Volatility*y, respectively. All market condition variables are standardized. In this analysis, we focus on retail investors' real trades only. We control for *Stock Market Beta*, which is the beta coefficient estimated by regressing the stock's excess weekly returns on the value-weighted excess market returns over the past 52 weeks.. We also include the log of *Total Stocks*, the log of *Trade Size*, and *Years of Experience*. Standard errors are clustered at the investor and date levels. t-statistics are in parentheses. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	(1)	(2)	(3)
Number of Biases	-0.0123***	-0.0128***	-0.0144***
	(-9.16)	(-9.04)	(-9.80)
Number of Biases × Market Condition	-0.0090***	0.0065***	-0.0142***
	(-3.35)	(3.55)	(-12.94)
Market Condition Variable	DY ratio	Past Year HEX Return	Past Year HEX Volatility
Controls	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes
Date FE	Yes	Yes	Yes
Adj. R^2	0.446	0.446	0.447
Obs.	309,812	309,839	308,212

Table 8: **The Prevalence of Biases Across Market Conditions**

The table examines the relationship between the prevalence of biases and market conditions using Poisson regression. The dependent variable, *Number of Biases*, is the total count of biases associated with a single trade. We use several market condition variables. From Columns 1 to 3, we use *DY ratio*, *Past Year HEX Return*, and *Past Year HEX Volatility*, respectively. All market condition variables are standardized. In this analysis, we focus on retail investors' real trades only. We control for *Stock Market Beta*, which is the beta coefficient estimated by regressing the stock's excess weekly returns on the value-weighted excess market returns over the past 52 weeks.. We also include the log of *Total Stocks*, the log of *Trade Size*, and *Years of Experience*. Standard errors are clustered at the investor and date levels. t-statistics are in parentheses. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	(1)	(2)	(3)
DY ratio	0.0374***		
	(7.37)		
Past Year HEX Return		-0.0098***	
		(-2.77)	
Past Year HEX Volatility			0.0232***
			(9.15)
Controls	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes
Pseudo R^2	0.031	0.031	0.031
Obs.	310,069	310,096	308,463

Table 9: **The Prevalence of Biases Across Market Conditions and Investor Characteristics**

The table examines the relationship between bias exposure and the market conditions using Poisson regression. The dependent variable, *Number of Biases*, is the total count of biases associated with a single trade. From Columns 1 to 3, we test investor characteristics such as *Female*, *Age*, and *Years of Experience*. *Female* is a binary indicator equal to 1 if the trade is made by a female trader and 0 if made by a male trader. *Age* is the trader's current age at the time of each trade. *Years of Experience* is the number of years since the investor's first trade to the current trade. From Columns 4 to 6, we add market condition variables such as *DY ratio*, *Past Year HEX Return*, and *Past Year HEX Volatility*, respectively. All market condition variables are standardized. With these market condition variables, we interact investor characteristics. In this analysis, we focus on retail investors' real trades only. We control for *Stock Market Beta*, which is the beta coefficient estimated by regressing the stock's excess weekly returns on the value-weighted excess market returns over the past 52 weeks.. We also include the log of *Trade Size* as a control variable. Standard errors are clustered at the investor and date levels. *t*-statistics are in parentheses. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.0134***					
	(3.73)					
Age	0.0009***	0.0273***				
	(8.83)	(7.90)				
Years of Experience	0.0080***	-0.0243***	0.0221***	0.0258***	0.0246***	0.0222***
	(14.16)	(-7.00)	(9.59)	(10.94)	(10.62)	(9.38)
Female×Market Condition				0.0025	-0.0053*	0.0039
				(0.77)	(-1.78)	(1.30)
Age×Market Condition				0.0002**	-0.0002**	-0.0002***
				(2.57)	(-2.00)	(-2.70)
Years of Experience×Market Condition				-0.0027***	0.0030***	-0.0013***
				(-6.90)	(6.18)	(-2.69)
Market Condition Variable				DY ratio	Past Year HEX Return	Past Year HEX Volatility
Controls	No	Yes	Yes	Yes	Yes	Yes
Investor FE	No	Yes	Yes	Yes	Yes	Yes
Date FE	No	No	Yes	Yes	Yes	Yes
Pseudo R^2	0.002	0.031	0.043	0.043	0.043	0.043
Obs.	362,415	309,535	309,535	309,508	309,535	307,904

Appendix A: Additional Figures and Tables

Figure A1: **Bootstrap selection results across 100 iterations**

A total of 100 random draws, with 10 papers being drawn each time from a pool of 100 manually screened papers, was performed to test the accuracy and robustness of our two-step determination process, as described in Subsection 2.2. The figure presents the number of times each paper within that 100-paper testing pool was selected during the 100 random draws. We ensure that each paper was selected at least twice.

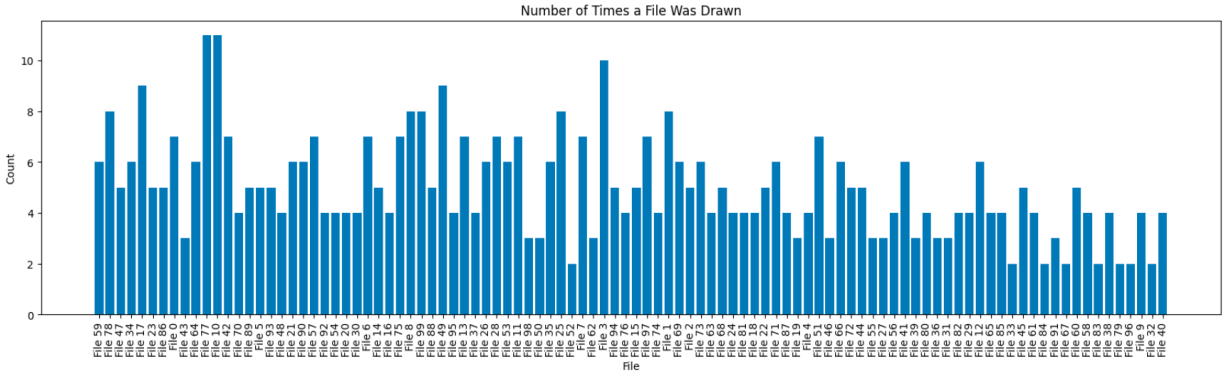


Figure A2: **Institution: The Relationship Between Individual Biases and Future Returns**

The figure shows the relationship between individual bias exposure and trading performance. The dependent variable, *Future Return*, represents the raw return of the stock traded over a specific time windows following the trade, measured over a 12-month window following the trade. The main independent variable, $Bias_b$, is defined as an indicator for each trade, which equals one if the trade meets the definition of bias b. In this figure, we include 18 biases commonly used by retail investors and highlight in bold the 6 biases that are also prevalent among institutional investors. Circle markers represent coefficient estimates obtained from a baseline regression of *Future Return* on $Bias_b$ without any controls, using real trading data exclusively. Square markers denote coefficient estimates from a regression of *Future Return* on $Bias_b$, including control variables and investor fixed effects, also using real trading data. Diamond markers indicate coefficient estimates for interaction terms between $Bias_b$ and *Real*, defined as 1 if the trade is from the real trading sample and 0 if from the simulations, from regressions incorporating investor-day fixed effects and using both real and simulated data. Standard errors are clustered at the investor and date levels. Control variables include *Stock Market Beta*, the logarithm of *Total Stocks*, the logarithm of *Trade Size*, and *Years of Experience*.

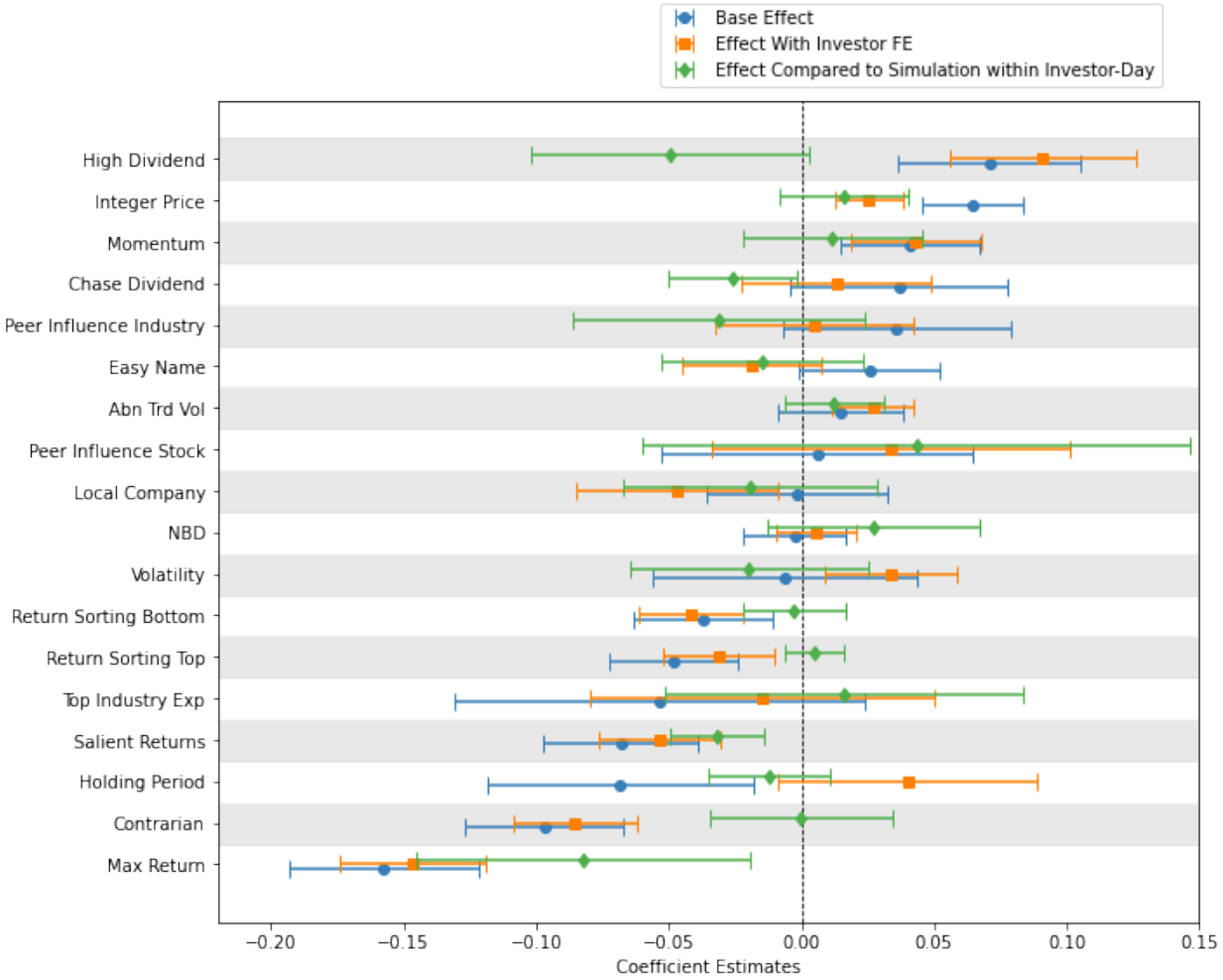


Figure A3: **The Relationship Between Individual Biases and Future Market-adjusted Returns**

The figure shows the relationship between individual bias exposure and trading performance. The dependent variable, *Future market-adjusted return*, represents the market-adjusted return of the stock traded over a specific time windows following the trade, i.e., stock excess return adjusted to stock market beta and excess return of HEX index, measured over a 12-month window following the trade. The main independent variable, $Bias_b$, is defined as an indicator for each trade, which equals one if the trade meets the definition of bias b. In this figure, we include 18 biases commonly used by retail investors. Circle markers represent coefficient estimates obtained from a baseline regression of *Future market-adjusted return* on $Bias_b$ without any controls, using real trading data exclusively. Square markers denote coefficient estimates from a regression of *Future market-adjusted return* on $Bias_b$, including control variables and investor fixed effects, also using real trading data. Diamond markers indicate coefficient estimates for interaction terms between $Bias_b$ and *Real*, defined as 1 if the trade is from the real trading sample and 0 if from the simulations, from regressions incorporating investor-day fixed effects and using both real and simulated data. Standard errors are clustered at the investor and date levels. Control variables include *Stock Market Beta*, the logarithm of *Total Stocks*, the logarithm of *Trade Size*, the logarithm of *Years of Experience*, and the logarithm of *Age*.

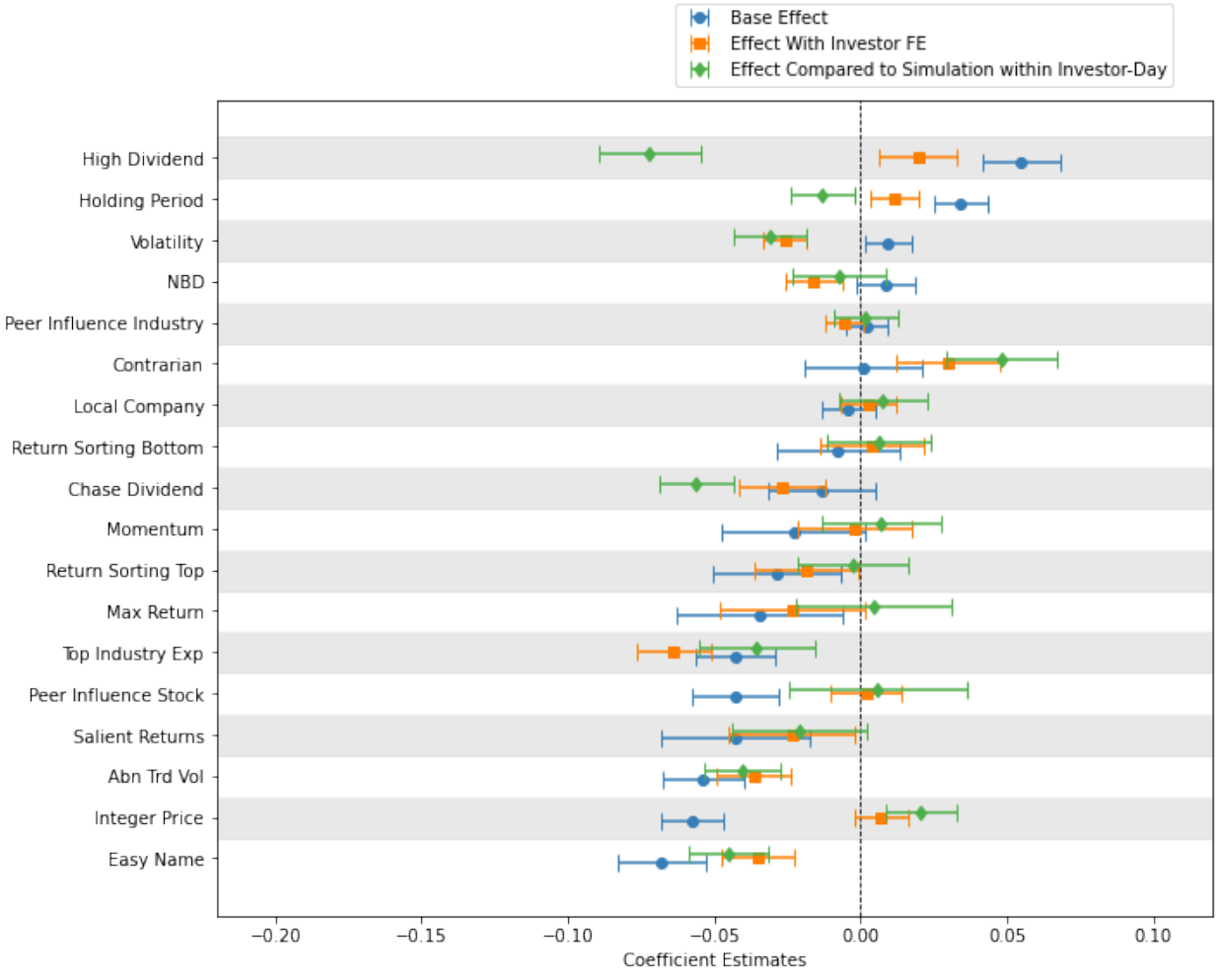


Table A1: **List of Biases**

This table presents the names, definitions, and references for the 22 selected biases, listed in alphabetical order.

Bias	**Definition**	**Reference**
Abnormal Trading Volume	Indicator equals 1 if the stock purchased has abnormal trading volume (Barber and Odean, 2008) above the median.	Barber and Odean (2008); Da et al. (2011); Gargano and Rossi (2018)
Chase Dividend	Indicator equals 1 if the stock is purchased on a day that falls between the stock's dividend announcement day and record date for dividend claim.	Rantapuska (2008)
Contrarian	Indicator equals 1 if the stock purchased has past 20-day return in the bottom 10 percentile.	Lakonishok et al. (1994); Grinblatt and Keloharju (2001b); Kaniel et al. (2008)
Easy Name	Indicator equals 1 if the stock name consists of a single word.	Green and Jame (2013)
Holding Period	Indicator equals 1 if the investor buys a stock they have held for at least 6 months.	Massa and Simonov (2006)
High Dividend	Indicator equals 1 if the stock purchased falls into the top decile of all stocks sorted on last year's dividends.	Graham and Kumar (2006); Hartzmark and Solomon (2019); Daniel et al. (2021)
Top Industry Experience	Indicator equals 1 if a stock purchased is not previously owned and in an industry where the investor experienced a top 10 percentile market-adjusted return in the previous year.	Kaustia and Knüpfer (2008); Huang (2019); Malmendier (2021)
Integer Price	Indicator equals 1 if the transaction price is within 5 cents distance to an integer price.	Kuo et al. (2015)
Local Company	Indicator equals 1 if the stock purchased has the same municipality as that of the investor.	Huberman (2001); Grinblatt and Keloharju (2001a); Feng and Seasholes (2004); Ivković and Weisbenner (2005); Seasholes and Zhu (2010)
Lottery-Like Stocks	Indicator equals 1 if the stock purchased is in the highest 50th percentile of previous 6 month idiosyncratic volatility, and in the highest 50th percentile of previous 6 month idiosyncratic skewness, and in the lowest 50th percentile of stock price.	Kumar (2009); Eraker and Ready (2015)

Table A1 continued...

Bias	**Definition**	**Reference**
Low Price	Indicator equals 1 if the stock purchased falls into the bottom decile of all stocks sorted on yesterday's nominal prices in the same currency (Euro).	Green and Hwang (2009); Birru and Wang (2016)
Max Return	Indicator equals 1 if the stock purchased falls into the top decile of all stocks sorted on their last month's maximum daily returns.	Bali et al. (2011); Liu et al. (2020)
Momentum	Indicator equals 1 if the stock purchased has past 20-day return in the top 10 percentile.	Grinblatt and Keloharju (2000); Liao et al. (2022)
Naive Buying Diversification	Indicator equals 1 if multiple stocks are purchased on the same day with the same amount of money up to a difference of 5%.	Gathergood et al. (2023)
Peer Influence Industry, Peer Influence Stock	Peer Influence Industry (Peer Influence Industry) equals 1 if the ratio of the fraction of investors in the same zip code who purchased a stock in the same industry (the same stock) to the fraction of all other investors who purchased stocks in the same industry (the same stock) is the top industry (the top 10% of stocks) within that zip-code month, and the ratio is at least 1.	Ivković and Weisbenner (2007); Hvide and Östberg (2015); Huang et al. (2021); Escobar and Pedraza (2023)
Return Sorting Top, Return Sorting Bottom	Return Sorting Top (Return Sorting Bottom) equals 1 if the stock purchased falls into top (bottom) 10 percentile of all stocks sorted on yesterday's returns.	Jegadeesh and Titman (1993); Grinblatt and Keloharju (2001b); Da et al. (2021); Barber et al. (2022)
Salient Returns	Indicator equals 1 if the stock purchased is in the top 10 percentile of salient returns (measured following Cosemans and Frehen (2021) over the past month.	Cosemans and Frehen (2021)
Small Cap	Indicator equals 1 if the stock purchased is in the bottom 10 percentile of all stocks sorted on prior day's market capitalization.	Kumar and Lee (2006)
Value	Indicator equals 1 if the stock purchased is in the top 10 percentile of all stocks sorted on book-to-market.	Kumar and Lee (2006)
Volatility	Indicator equals 1 if the stock purchased has a past year daily realized volatility that falls into an interval of 10% deviation from the dollar-weighted average realized volatility of other stocks in the portfolio.	Dorn and Huberman (2010)

Table A2: **The Relationship Between Stock Selection Biases and Performance – Future market-adjusted returns**

The table shows the relationship between the future return of a trade and bias exposure in the trade. The dependent variable, *Future market-adjusted return*, is the market-adjusted return of the stock traded over a specific time windows following the trade, i.e., stock excess return adjusted to stock market beta and excess return of HEX index. We use a 12-month window after the trade. The main independent variable, *Number of Biases*, is the total count of biases associated with a single trade. From Columns 1 to 3, we focus on the real trading data. In Column 4, we focus on the simulated trading data. In Column 5, we include the interaction terms of *Number of Biases* with *Real*, which is 1 if the trade is from the real trading sample and 0 if from the simulations. In all columns we control for *Stock Market Beta*, which is the beta coefficient estimated by regressing the stock's excess weekly returns on the value-weighted excess market returns over the past 52 weeks.. We also include the logarithm of *Total Stocks*, the logarithm of *Trade Size*, *Years of Experience*, and the logarithm of *Age* as control variables. Standard errors are clustered at the investor and date levels. t-statistics are in parentheses. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	(1)	(2)	(3)	(4)	(5)
Number of Biases	-0.0189***	-0.0165***	-0.0108***	0.0046***	0.0046***
	(-11.97)	(-10.16)	(-7.38)	(3.84)	(3.84)
Real × Number of Biases					-0.0154***
					(-9.37)
Data Subset	Real	Real	Real	Simulation	Both
Controls	No	Yes	Yes	Yes	Yes
Investor FEs	No	No	Yes	No	No
Investor×Subset FEs	No	No	No	Yes	Yes
Date FEs	No	No	No	No	No
Adj. R^2	0.0032	0.0290	0.1319	0.1619	0.1462
Obs.	312,794	312,736	306,624	282,841	589,465

Table A3: **The Prevalence of Biases Across Market Conditions – OLS regression**

The table replicates Table 8 using OLS regression. The dependent variable, *Number of Biases*, is the total count of biases associated with a single trade. We use several market condition variables. From Columns 1 to 3, we use *DY ratio*, *Past Year HEX Return*, and *Past Year HEX Volatility*, respectively. All market condition variables are standardized. In this analysis, we focus on retail investors' real trades only. We control for *Stock Market Beta*, which is the beta coefficient estimated by regressing the stock's excess weekly returns on the value-weighted excess market returns over the past 52 weeks.. We also include the log of *Total Stocks*, the log of *Trade Size*, and *Years of Experience*. Standard errors are clustered at the investor and date levels. *t*-statistics are in parentheses. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	(1)	(2)	(3)
DY ratio	0.1418***		
	(7.15)		
Past Year HEX Return		-0.0368***	
		(-2.76)	
Past Year HEX Volatility			0.0885***
			(9.08)
Controls	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes
Adj. R^2	0.112	0.110	0.112 Obs.
310,073	310,100	308,467	

Table A4: **The Prevalence of Biases Across Market Conditions and Investor Characteristics – OLS regression**

The table replicates Table 9 using OLS regression.The dependent variable, *Number of Biases*, is the total count of biases associated with a single trade. From Columns 1 to 3, we test investor characteristics such as *Female*, *Age*, and *Years of Experience*. *Female* is a binary indicator equal to 1 if the trade is made by a female trader and 0 if made by a male trader. *Age* is the trader's current age at the time of each trade. *Years of Experience* is the number of years since the investor's first trade to the current trade. From Columns 4 to 6, we add market condition variables such as *DY ratio*, *Past Year HEX Return*, and *Past Year HEX Volatility*, respectively. All market condition variables are standardized. With these market condition variables, we interact investor characteristics. In this analysis, we focus on retail investors' real trades only. We control for *Stock Market Beta*, which is the beta coefficient estimated by regressing the stock's excess weekly returns on the value-weighted excess market returns over the past 52 weeks.. We also include the log of *Trade Size* as a control variable. Standard errors are clustered at the investor and date levels. t-statistics are in parentheses. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.0487***					
	(3.73)					
Age	0.0032***	0.0963***				
	(8.81)	(7.88)				
Years of Experience	0.0289***	-0.0856***	0.0803***	0.0915***	0.0884***	0.0811***
	(14.14)	(-6.95)	(9.74)	(10.80)	(10.64)	(9.46)
Female×Market Condition				0.0118	-0.0201*	0.0136
				(0.90)	(-1.73)	(1.21)
Age×Market Condition				0.0010***	-0.0007**	-0.0008**
				(2.82)	(-2.21)	(-2.49)
Years of Experience×Market Condition				-0.0088***	0.0100***	-0.0040**
				(-5.75)	(5.49)	(-2.20)
Market Condition Variable				DY ratio	Past Year HEX Return	Past Year HEX Volatility
Controls	No	Yes	Yes	Yes	Yes	Yes
Investor FE	No	Yes	Yes	Yes	Yes	Yes
Date FE	No	No	Yes	Yes	Yes	Yes
Adj. R^2	0.008	0.110	0.176	0.176	0.176	0.177
Obs.	362,415	309,539	309,539	309,512	309,539	307,908